INTERNATIONAL STANDARD

ISO 8253-3

Second edition
2012-03-01

Acoustics — Audiometric test methods —

Part 3: Speech audiometry

Acoustique — Méthodes d'essais audiométriques —

Partie 3: Audiométrie vocale

Reference number
ISO 8253-3:2012(E)

COPYRIGHT PROTECTED DOCUMENT

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

ISO 8253-3 was prepared by Technical Committee ISO/TC 43, *Acoustics*.

This second edition cancels and replaces the first edition (ISO 8253-3:1996), which has been technically revised.

ISO 8253 consists of the following parts, under the general title *Acoustics — Audiometric test methods*:

— *Part 1: Pure-tone air and bone conduction threshold audiometry*

— *Part 2: Sound field audiometry with pure-tone and narrow-band test signals*

— *Part 3: Speech audiometry*

Introduction

Speech audiometry is used for the assessment of hearing in connection with diagnostic evaluation and audiological rehabilitation.

The results of speech audiometry depend on the speech material and test method used. This part of ISO 8253 sets conditions for speech materials in order to assure minimum requirements of precision and comparability between different tests using different speech materials including materials in different languages. It also specifies procedures to be used when testing speech recognition.

Acoustics — Audiometric test methods —

Part 3:
Speech audiometry

1 Scope

This part of ISO 8253 specifies basic methods for speech recognition tests for audiological applications.

In order to ensure minimum requirements of precision and comparability between different test procedures including speech recognition tests in different languages, this part of ISO 8253 specifies requirements for the composition, validation and evaluation of speech test materials, and the realization of speech recognition tests. This part of ISO 8253 does not specify the contents of the speech material because of the variety of languages.

Furthermore, this part of ISO 8253 also specifies the determination of reference values and fulfilment requirements for the realization and manner of presentation.

This part of ISO 8253 specifies procedures and requirements for speech audiometry with the recorded test material being presented by air conduction through an earphone, or from a loudspeaker for sound field audiometry. Methods for using noise either for masking the non-test ear or as a competing sound are described.

Some test subjects, for example children, can require amended test procedures not specified in this part of ISO 8253.

Specialized tests such as those used for evaluating directional hearing and dichotic hearing are outside the scope of this part of ISO 8253.

2 Normative references

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

ISO 266, *Acoustics — Preferred frequencies*

ISO 8253-1, *Acoustics — Audiometric test methods — Part 1: Pure-tone air and bone conduction audiometry*

ISO 8253-2, *Acoustics — Audiometric test methods — Part 2: Sound field audiometry with pure-tone and narrow-band test signals*

ISO/IEC Guide 98-3, *Uncertainty of measurement — Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*

IEC 60645-1, *Electroacoustics — Audiological equipment — Part 1: Pure-tone audiometers*

IEC 60645-2:1993, *Audiometers — Part 2: Equipment for speech audiometry*

IEC 61672-1, *Electroacoustics — Sound level meters — Part 1: Specifications*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 8253-1 and ISO 8253-2 and the following apply.

3.1
speech signal
acoustic signal which carries information in a given language

NOTE A speech signal may be a voice signal or an acoustic signal simulating a voice signal.

3.2
test item
particular monosyllabic or polysyllabic word or logatom, or sentence, or time-limited segment of connected speech, used in accordance with defined rules of presentation and scoring in a speech audiometric procedure

NOTE Scoring may be based on a complete test item or parts thereof being correctly recognized.

3.3
speech material
entire set of test items which is used for speech recognition tests

NOTE Usually the speech material is subdivided into several test lists.

3.4
open-set test material
set of test items in which the number of alternative responses to each test item is unlimited

3.5
closed-set test material
set of test items in which the number of alternative responses to each test item is limited

3.6
phoneme
smallest recognizable unit of speech in a given natural language

3.7
phoneme class
subdivision of phonemes that show characteristic similarities in vocal production mode as well as in acoustical signal properties

3.8
syllable
segment of speech which consists of a vowel with or without one or more accompanying consonants immediately preceding or following

3.9
spondee
two-syllable word with equal stress on each syllable

3.10
logatom
syllabic unit that has no verbal meaning to the listener

NOTE A logatom is sometimes called a “nonsense syllable”.

3.11
test list
number of selected test items, presented and scored as a single unit

3.12
set of test items
selected number of test items from a test list

3.13
phonemically balanced test list
test list which contains approximately the same proportions of the various phonemes as are present in typical speech communication within a given language

3.14
perceptually balanced test list
test list which yield equivalent speech recognition score with minimized variance at a specified speech level under equivalent conditions (specified manner of presentation and specified group of listeners)

NOTE Usually the specified group of listeners consists of otologically normal persons with an age from 18 years to 25 years inclusive. For special tests, different groups (e.g. otologically normal children within a specified age range) may be used as well.

3.15
carrier phrase
sentence or phrase in which a test item is embedded such that the correct recognition of the test item is not dependent upon the context or meaning of the sentence or phrase

3.16
reference recording of speech material
master recording
well-defined recording that represents the speech material and that is used for validation and application of the speech material

3.17
speech level
equivalent continuous sound pressure level of the speech material as measured in an appropriate coupler, ear simulator or in a sound field, with frequency weighting C according to IEC 61672-1

NOTE For speech test lists based on single test items separated by silent intervals, the integration does not include these intervals. For test lists based on single test items with a carrier phrase, the integration includes the test items only. In a digital recording, the silent intervals can be removed by editing. Numerical correction may also be made by determination of the total duration of the test items and the total duration of the silent intervals.

3.18
equivalent continuous sound pressure level
$L_{p,\mathrm{eq}T}$
ten times the logarithm to the base 10 of the ratio of the time average of the square of the sound pressure, p, during a stated time interval of duration, T (starting at t_1 and ending at t_2), to the square of a reference value, p_0, expressed in decibels

$$L_{p,\mathrm{eq}T} = 10 \lg \left[\frac{\frac{1}{\mathrm{T}} \int_{t_1}^{t_2} p^2(t)\,\mathrm{d}t}{p_0^2} \right] \mathrm{dB}$$

where the reference value, p_0, is 20 μPa

NOTE Because of practical limitations of the measuring instruments, p^2 is always understood to denote the square of a frequency-weighted and frequency-band-limited sound pressure. If a specific frequency weighting as specified in IEC 61672-1 and/or specific frequency bands are applied, this should be indicated by appropriate subscripts; e.g. $L_{p,\mathrm{A},10\ \mathrm{s}}$ denotes the A-weighted time-averaged sound pressure level over 10 s.

[ISO/TR 25417[3], 2.3]

3.19
speech detection threshold level
for a given test subject, for specified speech material and a specified manner of signal presentation, the speech level at which it is detected (not necessarily understood) in 50 % of the trials

3.20
speech recognition score
for a given test subject, for specified speech material, a specified manner of signal presentation and at a specified speech level, the percentage of correctly recognized test items or scorable items if the scoring method is not based on whole test items

3.21
maximum speech recognition score
for a given test subject, for specified speech material and a specified manner of signal presentation, the maximum speech recognition score obtained regardless of speech level

3.22
speech recognition threshold level
for a given test subject, for specified speech material and a specified manner of signal presentation, the lowest speech level or speech-to-noise ratio at which the speech recognition score is equal to 50 %

NOTE Speech recognition threshold has been called “speech reception threshold”.

3.23
reference speech recognition threshold level
for specified speech material and a specified manner of signal presentation, the median value of the speech recognition threshold levels of a sufficiently large number of otologically normal persons, of both sexes, between 18 years and 25 years inclusive for whom the test material is appropriate

3.24
optimum speech level
for a given test subject, for specified speech material and a specified manner of signal presentation, the speech level or range of speech levels at which maximum speech recognition score is obtained

3.25
half-optimum speech level
for a given test subject, for specified speech material and a specified manner of signal presentation, the speech level at which half of the maximum speech recognition score is obtained

3.26
speech recognition curve
for a given test subject, for specified speech material and a specified manner of signal presentation, a curve that describes the speech recognition score as a function of speech level

3.27
reference speech recognition curve
for specified speech material and a specified manner of presentation, a curve that describes the median speech recognition score as a function of speech level for a sufficiently large number of otologically normal persons of both sexes, aged between 18 years and 25 years inclusive and for whom the test material is appropriate

3.28
hearing level for speech
for specified speech material and a specified manner of signal presentation, the speech level minus the appropriate reference speech recognition threshold level

3.29
competing sound
additional sound that is presented during speech recognition tests

3.30
competing sound pressure level
equivalent continuous sound pressure level of a competing sound as measured in an appropriate coupler, ear simulator or in a sound field, with frequency weighting C according to IEC 61672-1

3.31
speech-to-noise level difference
speech-to-noise ratio
SNR
difference between the speech level and the competing sound level

3.32
effective masking level for speech
level of a specified masking sound which is numerically equal to that hearing level for speech to which the speech recognition threshold level for a specified speech signal would be raised by the presence of that masking sound

3.33
most comfortable level for speech
for a given test subject and a specified manner of presentation, the speech level at which the loudness of the speech signal is judged by the test subject to be the most comfortable

3.34
connected speech
running speech
continuous speech with natural intonation, consisting of consecutive sentences with logical connections

4 Requirements for recording of speech material

4.1 General requirements

Each copy of the recording shall contain the following signals:

a) the speech material;

b) a signal for the calibration of the equipment for speech audiometry, of a duration not less than 60 s; the calibration signal shall be a weighted random noise, e.g. as specified in IEC 60645-2:1993, 13.1, a band of noise centred at 1 kHz and having a bandwidth of one-third octave, or a frequency-modulated tone at 1 kHz having a bandwidth of at least one-third octave — the modulating signal shall be either sinusoidal or triangular with a repetition rate in the range from 4 Hz to 20 Hz;

c) signals for testing the frequency response of the equipment to be used for speech audiometry including the playback equipment and the recording — such signals shall have a duration of not less than 15 s and consist of one-third-octave bands of white noise centred at the preferred one-third-octave frequencies in accordance with ISO 266 in the frequency range 125 Hz to 8 000 Hz.

NOTE For speech material based on other time and/or frequency weightings than C-weighted equivalent sound pressure level, correction factors can be determined describing the difference related to C-weighted equivalent sound pressure level for speech level, calibration signal level and competing noise level.

4.2 Reference recording

A suitable speaker, either male or female, should have a normal and clear articulation without pronounced dialect. The speaker should be instructed to maintain clarity, natural pace, and vocal effort and avoid emphasis on key words.

A reference recording mainly consists of the elements specified in 4.1. The reference recording shall not be modified. All copies shall be made from the reference recording. Any changes of the reference recording (e.g. a new recording or changes of levels or frequency content) require a new validation (see Clause 5).

For speech audiometry intended to be performed in noise the reference recording shall also include the associated competing noise or noises, either on the same or on separate channels.

4.3 Recording environment

The room in which the speech is recorded shall be sufficiently quiet to provide a signal-to-noise ratio of at least 40 dB and shall have a reverberation time of less than 0,5 s at any frequency in the range from 125 Hz to 8 000 Hz.

4.4 Frequency response of recording equipment

The frequency response of the recording equipment including the microphone used shall be flat within 2 dB in the frequency range from 125 Hz to 8 000 Hz.

NOTE The quality of the recording and negligible loss of quality after repeated use make digital recordings superior.

4.5 Interval between successive test items

The intervals between the successive test items of the recording should be specified and constant within ±10 %.

For single test words, an interval of not less than 4 s is suggested.

4.6 Levels

4.6.1 Speech material

The average reference speech recognition threshold level of the smallest set of test items to be used at a predetermined sound pressure level shall not deviate by more than ±1 dB from the average of all test items within the list. The reference speech recognition threshold level of any single test item shall not deviate by more than ±3 dB from the average reference speech recognition threshold level of all test items of the list. The average speech recognition threshold level of test lists shall not differ from the average reference speech recognition threshold level of all test items recorded by more than ±1 dB, see Annex D.

4.6.2 Calibration signal

The level of the calibration signal shall not deviate by more than ±0,5 dB from the average speech level of all test items of the specific speech material. The level of the calibration signal shall be measured using the same frequency weighting as when measuring the speech level.

4.7 Signal-to-noise ratio

The level of any inherent background noise on the recording shall be at least 40 dB below the level of the calibration signal in accordance with 4.6.2 as measured using the same frequency weighting as when measuring the speech level.

4.8 Phonemic balance across test lists

All test lists of a specific speech material shall be phonemically balanced.

In some cases, it may not be possible to achieve perfect phonemic balance. In such cases, the test lists shall be phonemically balanced based on phoneme classes, i.e. voiced and unvoiced plosives and fricatives, nasals, long and short vowels.

The distribution of phonemes of the speech material should approximate the phoneme distribution of the test language as closely as possible. Speech materials consisting of a large amount of items allow a closer approximation than smaller speech materials. In some cases, the possibility of achieving a close approximation to the phoneme distribution is limited, e.g. in screening tests or tests for children, because the speech material has to be limited to a relatively small number of test items.

4.9 Perceptual balance across test lists

All test lists of a specific speech material shall be perceptually balanced, i.e. the result of the speech recognition test shall be independent of the choice of test list.

Perceptually balanced test lists shall fulfil the requirements given in 4.6.1. Additional requirements for perceptually balanced test lists are given in the following.

The speech recognition curves for all test lists shall have been determined in such a way that the speech recognition threshold level can be determined with an expanded uncertainty not exceeding 1 dB (the width of the 95 % confidence interval shall not exceed 2 dB); see Annex E. To achieve this, speech recognition measurements at different presentation levels (or speech-to-noise ratios in case of a speech recognition test in noise) shall be performed with at least 10 otologically normal persons whose native language is the same as the language of the speech material. The speech recognition measurements shall be performed within the normal context of the speech stimulus (e.g. presentation of sentences when using a sentence test). The speech recognition of each particular test item shall then be scored and analysed independently (e.g. scoring of particular words of the sentences).

In order to achieve an expanded uncertainty not exceeding 1 dB, extreme care is necessary to reduce the potential uncertainty sources.

In order to increase the perceptual homogeneity of the speech material, speech recognition differences between the particular speech items may be minimized (see Annex D).

5 Validation of speech material recordings

5.1 General

For the validation, equipment fulfilling the requirements of IEC 60645-2:1993, Clauses 6 to 17, type A, A-E, B or B-E shall be used. Testing shall be performed monaurally. If equipment of type A-E or B-E is used, the evaluation shall include using both uncorrected earphone and earphone corrected for free-field equivalent output. Only complete test lists shall be used.

NOTE It can be expected that the use of equipment of Type A-E or B-E gives identical results for both binaural earphone and loudspeaker listening due to their free-field equivalent frequency response.

5.2 Determination of reference speech recognition curve

The reference speech recognition curve for each speech material and manner of presentation shall be determined by performing speech recognition tests using a sufficiently large group (at least 25) of otologically normal persons whose first language is the language of the speech material. The minimum requirements for the pure-tone hearing threshold levels of test subjects shall be specified by the laboratory carrying out the validation.

It is recommended that the following criterion based on the pure-tone audiogram be used: the hearing threshold levels of the test subjects should be at or below 10 dB at the frequencies 250 Hz, 500 Hz, 750 Hz, 1 000 Hz, 1 500 Hz, 2 000 Hz, 3 000 Hz, 4 000 Hz, 6 000 Hz, and 8 000 Hz. At a maximum of two frequencies the hearing threshold level may be 15 dB.

If the speech recognition test is intended to be performed with children, the reference speech recognition curve shall be determined using a sufficiently large group (at least 25) of otologically normal children of the same age range as the speech material is intended to be used for.

If the speech recognition test is intended to be used in silence, the reference speech recognition curve shall be determined in silence.

If the speech recognition test is intended to be used with a competing noise, the reference speech recognition curve shall be determined using this noise. If the test is intended to be used with different noises, the reference speech recognition curve shall determined using each particular noise.

The reference speech recognition curve shall be given as a table which describes the relation between speech level (or speech-to-noise level difference) and speech recognition score. The speech levels (or speech-to-noise level differences) which are required to achieve the speech recognition scores 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, and 90 % shall be specified. These values can be calculated using an adequate interpolation between the measured data.

If persons show a large variability of speech recognition results (e.g. children), it can prove necessary to measure the reference speech recognition curve using a larger number of persons.

If it is difficult to determine the whole reference speech recognition curve from 0 % to 100 %, ensure that the speech recognition curve is determined in the range of interest. If the speech recognition test is intended to be used for the determination of speech recognition threshold levels, the reference speech recognition curve should be measured at values near the speech recognition threshold level (including values above and below this value) using a sufficient number of measurements. If the speech recognition test is intended to be used for the determination of the optimum speech level, the reference speech recognition curve should be measured at values from 50 % (or less) to values near to 100 % using a sufficient number of measurements.

5.3 Determination of perceptual equivalence of test lists

The perceptual equivalence of test lists of speech materials fulfilling the requirements specified in 4.9 shall be determined by performing speech recognition tests using a sufficiently large group (at least 10) of otologically normal persons as described in 5.2.

If the speech recognition test is intended to be used in silence, the perceptual equivalence of test lists shall be determined in silence.

If the speech recognition test is intended to be performed with a competing noise, the perceptual equivalence of test lists shall be determined using this noise. If the test is intended to be performed with different noises, the perceptual equivalence of test lists shall be determined using each particular noise.

If persons show a large variability of speech recognition results (e.g. children), it may be necessary to determine the perceptual equivalence of test lists curve using a larger number of persons.

The perceptual equivalence of test lists shall be specified by the average 95 % confidence intervals of the test lists. If the speech recognition test is intended for the determination of speech recognition threshold levels, the equality of test lists shall be specified as the average 95 % confidence interval for the test lists of the speech recognition threshold level when using different test lists with the same group of test subjects. If the speech recognition test is intended for the determination of the optimum speech level, the equality of test lists shall be specified as the average 95 % confidence intervals for the test lists of the speech recognition scores 50 %, 60 %, 70 %, 80 %, and 90 % when using different test lists with the same group of test subjects.

That means that the squared differences between the results using the different test lists should be averaged across listeners. The square root of this value multiplied by 2 gives the desired value.

5.4 Determination of repeatability of results (average test-retest reliability)

The repeatability of a speech recognition test result is quantified by the average test-retest reliability within one subject. Generally, different test lists shall be used in test and retest because the responses of the initial test may be memorized by the subject which may influence the result of the retest.

If the speech recognition test is intended to be performed in silence, the average test-retest reliability shall be determined in silence.

If the speech recognition test is intended to be performed with a competing noise, the average test-retest reliability shall be determined using this noise. If the speech recognition test is intended to be performed with different noises, the average test-retest reliability shall be determined using each particular noise.

The average reliability (test versus retest) shall be specified by the average 95 % confidence intervals across subjects.

That means that the squared differences between the results of the test and the retest measurements should be averaged across listeners. The square root of this value multiplied by 2 gives the desired value.

If the speech recognition test is intended for the determination of speech recognition threshold levels, the test-retest reliability shall be assessed for the speech recognition threshold level. If the speech recognition test is intended for the determination of the optimum speech level, the test-retest reliability shall be specified for the speech recognition scores 50 %, 60 %, 70 %, 80 %, and 90 %.

5.5 Documentation

Each copy of the recording shall be supplied with documentation containing the following information:

a) a printed version of the recorded speech material;

b) the method(s) of scoring, e.g. by sentence, word, key word, or phoneme;

c) the reference speech recognition curves as specified in 5.2 for each method of scoring, the origin of the reference curves and details of measuring conditions (e.g. number, age range and range of hearing threshold levels of test subjects and type of equipment for speech audiometry used as well as the arrangement of loudspeakers for competing noise, if applicable);

d) the repeatability of results as specified in 5.4;

e) the durations, frequency spectra and levels relative to speech level of the calibration signal and other recorded signals for the purpose of testing performance specifications in accordance with 4.1 b) and c);

f) the level of the calibration signal and its relationship to the reference speech recognition threshold level;

g) the intervals between successive test items;

h) the characteristics of any accompanying competing noise signal.

6 Requirements for speech audiometry

Speech audiometry shall be performed using equipment fulfilling the requirements of IEC 60645-2. Testing may be performed either by earphones monaurally or binaurally, by bone vibrator, or in a sound field. The level of a speech signal shall be expressed in terms of either speech level or hearing level for speech for the particular test material being used.

7 Ambient sound pressure levels in test room for speech audiometry

The ambient sound pressure levels in the test room shall not mask the speech signals. The requirements for the ambient sound pressure levels in the test room depend on the mode of presentation of the speech signal, i.e. through an earphone, bone vibrator or loudspeaker.

NOTE Requirements for ambient noise during speech audiometry can be less stringent than those for pure-tone threshold audiometry. If a test room is appropriate for pure-tone threshold audiometry down to a certain hearing level value over the whole frequency range, the room can also serve for speech audiometry for the same mode of presentation of test signals for signal levels down to the same hearing level values. Maximum permissible ambient sound pressure levels for pure-tone audiometry in the frequency range from 125 Hz to 8 000 Hz, are specified in ISO 8253-1 using earphones and bone vibrators and in ISO 8253-2 for sound field audiometry.

8 Sound field speech audiometry

Sound field speech audiometry can be performed with and without competing sound reproduced especially for the purpose of the test.

The recommended position of the speech loudspeaker is frontal to the listener at an angle of incidence of 0°. If the speech audiometric results from an individual test subject are to be compared with the reference speech recognition threshold level or the reference speech recognition curve, the position of the speech loudspeaker shall be identical to that used for the determination of the reference values.

If noise is to be presented as a competing sound to the speech signal, the recommended positions of the noise loudspeakers are given in 15.2.

The type of sound field (free, quasi-free or diffuse), as specified in ISO 8253-2, and the position of the reference point shall be recorded.

9 Preparation and instruction of test subject

9.1 General

It is generally assumed that pure-tone audiometry has been performed prior to speech audiometry.

9.2 Preparation of test subject

Recent exposure to noise may cause a temporary elevation of the hearing threshold levels. Therefore, significant noise exposure should be avoided before audiometric testing or it shall be noted. In order to avoid errors due to excessive physical exertion, the subject should be present at least 5 min prior to testing.

Normally, the audiometric test is preceded by an otoscopic examination carried out by a qualified person. If obstructing wax is found in the ear canals, remove it and delay audiometry for a suitable period. Also check the ears for the possibility of collapsing ear canals when earphones are to be used and take appropriate action.

NOTE It is possible that requirements for a qualified person are specified by national authorities or other suitable organizations.

The ability of the test subject to understand the test material shall be established prior to testing. If the response to the test material is to be spoken, it shall be established that the test subject can reproduce the test material.

9.3 Instruction of test subject

In order to achieve reliable results, it is essential that relevant instructions in the test procedure be given unambiguously and that they are fully understood by the test subject.

The instructions shall be phrased in language appropriate to the test subject. They depend upon the type of speech test to be carried out. Instructions shall normally indicate:

a) which ear shall be tested first;

b) the type of test items and the response task;

c) the need to respond whenever the speech is heard in either ear, no matter how faint it may be;

d) the need to respond as soon as each test item has been heard;

e) if the response is spoken, to repeat clearly the test item heard, even when not certain that the test item has been heard correctly; only one response to each test item is permitted.

Test subjects shall also be instructed to avoid unnecessary movements so as to obviate extraneous noise. After the instructions have been given, the tester shall establish that they have been understood. If there is any doubt, the instructions should be repeated. The test subject shall be informed that the test may be interrupted in case of any discomfort.

10 Subject's response mode

The test subject's response is normally spoken but may be written or indicated by means of a keyboard. When the tester is not in the same room as the test subject and the test subject's response to the speech signal is spoken, a talk-back system shall be used.

The spoken response of the test subject shall be clearly understandable by the tester. If not, an alternative response mode shall be used. The tester shall be familiar with the language and have hearing appropriate for understanding the spoken responses of the test subject.

The interval between the test items depends upon the speech test material, the purpose of the speech test and the subject's response mode and shall be stated.

The speech recognition score of a test subject may depend on the intervals between the test items. To make test results comparable to corresponding reference values, use intervals identical to those employed during the determination of reference values.

11 Determination of speech detection threshold level

11.1 In monaural testing, start with the ear considered to be more sensitive.

11.2 When determining speech detection threshold, connected speech is preferred as the speech signal. Let each test item be a segment of 1 s to 2 s duration, controlled by means of the interrupter switch of the equipment. The first item shall be presented at a sufficiently high level to evoke a definite response; e.g. 30 dB above the average of the subject's pure-tone hearing threshold levels at 500 Hz, 1 000 Hz, and 2 000 Hz.

11.3 Reduce the level in steps of 20 dB until the subject does not respond.

11.4 Increase the level of the speech signal in steps of 5 dB and present one test item on each level until the subject responds.

11.5 After the response, decrease the level by 10 dB and then begin another ascent using steps of 5 dB until the subject responds.

11.6 Repeat step 11.5 until three responses have occurred at the same level out of a maximum of five ascents. This level is the speech detection threshold level.

11.7 If less than three responses out of five ascents have been obtained at the same level, present a test item at a level 10 dB higher than the level of the last response. Then repeat steps 11.4 to 11.6 inclusive.

11.8 Proceed with the other ear if required.

11.9 A shortened version of this ascending method may be assumed to yield nearly equivalent results and may be appropriate in some cases. In this shortened version, continue the testing according to step 11.6 until at least two responses occur at the same level out of a maximum of three ascents.

12 Determination of speech recognition threshold level

12.1 General

12.1.1 Complete test lists shall be used. A test item may be a single word, a phrase or a sentence. A sentence may be scored as completely correct or by correctly repeated key words.

The descending procedures described in 12.2 and 12.3 are based on two widely used step-sizes: 5 dB and 2 dB. Two alternative procedures are described:

a) a descending procedure using 5 dB steps;

b) an alternative descending procedure using 2 dB or 5 dB steps.

The procedures are expected to yield comparable results. However, experimental evidence for this is unavailable at the time of publication. The procedures are identical from 12.1.2 to 12.1.4. Then, the descending procedure using 5 dB steps continues with 12.2 and the alternative procedure continues with 12.3.

In 12.4 and 12.5 two adaptive test methods are described, one based on fixed step sizes and the other on varying step sizes.

12.1.2 In monaural testing, start with the ear considered to be more sensitive.

12.1.3 The test subject shall be familiarized with the task prior to threshold determination by the presentation of a number of test items at a sufficiently high level to be clearly audible. A hearing level of speech of 20 dB to 30 dB above the average of the subject's pure-tone hearing threshold levels at 500 Hz, 1 000 Hz, and 2 000 Hz is generally recommended.

12.1.4 Reduce the speech level in steps of 5 dB, presenting at least two test items on each level, until the test subject no longer responds correctly to all test items. Continue according to 12.2 or 12.3.

12.2 Descending procedure using 5 dB steps

12.2.1 Present a set of test items at the level where the subject ceased to respond correctly according to 12.1.3 and record the number of correct responses. A set of test items shall contain at least 10 items.

12.2.2 If the subject scores at least 50 % on the set of test items, reduce the level in steps of 5 dB and present a new set of test items on each level until the subject scores less than 50 % on the set of test items. Usually one level is found to yield somewhat more than 50 % and the next lower level somewhat less than 50 %. If the subject scores less than 50 % on the set, increase the level in steps of 5 dB and present a new set of test items on each level until the subject scores more than 50 % on the set of test items.

12.2.3 The speech recognition threshold level for the ear is the integer value of the level corresponding to a score of 50 % correct as calculated by means of linear interpolation between the lowest level that yielded more than 50 % correct responses and the highest level that yielded less than 50 % correct responses. If a score of exactly 50 % is obtained at one speech level, that level represents the speech recognition threshold.

12.3 Alternative descending procedure

12.3.1 When one test item is missed, present a second at the same level. Continue this descending process in steps of 10 dB until a level is reached at which two consecutive test items are missed at the same level.

12.3.2 Increase the speech level by 10 dB. This defines the starting speech level, L_s.

12.3.3 Present two test items at this starting speech level and at each successive 2 dB decrement if a step size of 2 dB is used. For a step size of 5 dB, see step 12.3.6.

12.3.4 Continue this process if at least five out of the first six test items are repeated correctly. If this criterion is not met, increase the starting speech level by 4 dB to 10 dB.

12.3.5 The descending series is terminated when the test subject responds incorrectly to five of the last six test items presented.

12.3.6 If a step size of 5 dB is used, five test items should be presented at each level and the test terminated when all test items at a single level are incorrectly recognized.

12.3.7 Calculate the speech recognition threshold level, T, in decibels, according to the following equation:

$$T = L_S - \frac{dr}{n} + \frac{d}{2}$$

where

L_S is the starting speech level, in decibels, according to step 12.3.2;

r is the total number of correct responses given from the starting speech level according to step 12.3.3 up to the termination according to step 12.3.5 or 12.3.6;

d is the step size, in decibels;

n is the number of test items per level.

12.4 Adaptive procedure using fixed step sizes

12.4.1 The following is an example of an adaptive procedure using fixed step size based on identifying the speech level where 50 % of the test items are correctly recognized. When alternative methods are used, e.g. in order to determine a speech level that corresponds to a percentage of correctly identified test items other than 50 %, the differences shall be clearly specified.

12.4.2 In monaural testing, start with the ear considered to be more sensitive.

12.4.3 The test subject shall be familiarized with the task prior to threshold determination by the presentation of test items at a sufficiently high level to be clearly audible. A hearing level of speech of 20 dB to 30 dB above the average of the subject's pure-tone hearing threshold levels at 500 Hz, 1 000 Hz, and 2 000 Hz is generally recommended.

12.4.4 Reduce the speech level in steps of 5 dB, presenting one test item on each level, until the test subject no longer responds correctly to a test item. Continue according to 12.4.5 or 12.5.

12.4.5 Decrease the level by 2 dB and present the first test item of a new test list.

12.4.6 Continue in accordance with 12.4.7 to 12.4.9 for single words or complete sentences as test items. For scoring based on key words in sentences, follow steps 12.4.10 to 12.4.12.

12.4.7 If the listener is able to repeat the item correctly, decrease the level by 2 dB. If not, increase the level by 2 dB. Present another test item.

12.4.8 Repeat 12.4.6 for all remaining test items of the test list.

12.4.9 Discarding the first two test items, average the presentation levels of the remaining items of the test list. This average is the speech recognition threshold.

12.4.10 If the listener is able to repeat correctly more than half of the key words, decrease the level by 2 dB. If the number of correctly repeated key words is less than half, increase the level by 2 dB. If the number of correctly repeated key words is exactly half, leave the speech level unchanged and present another test sentence.

12.4.11 Repeat 12.4.10 for all remaining sentences of the test list.

12.4.12 Discarding the first two sentences, average the presentation levels of the remaining sentences of the test list. This average is the speech recognition threshold.

12.5 Other adaptive procedure

It may be reasonable to use a different adaptive rule than the one described in 12.4. In such a case, reference values as well as the test-retest reliability of the adaptive procedure in combination with the speech material shall be determined using a sufficiently large group of otologically normal persons as described in 5.1. The settings of the adaptive procedure (step sizes, word scoring or sentence scoring, target intelligibility and any other variations from the procedure described in 12.4) have to be specified.

13 Determination of speech recognition scores

If speech recognition scores are to be determined at more than one level (e.g. to determine a speech recognition curve), test items should not be repeated within the same session. The score shall always be based on the presentation of a complete test list.

If a test is not preceded by a determination of speech recognition threshold, the test subject shall be familiarized with the task prior to the determination of a score by the presentation of a number of test items at a sufficiently high level to be clearly audible. Usually, a hearing level for speech of 30 dB to 40 dB above the subject's average pure-tone hearing threshold levels at 500 Hz, 1 000 Hz, and 2 000 Hz is a good choice.

Set the audiometric equipment to the speech level required and present a complete test list at each level.

The choice of test level(s) is governed by the particular purpose of the test, as follows.

a) To determine a maximum speech recognition score, a first score is normally determined at a speech level of 25 dB or 30 dB above the speech recognition threshold level. The level should then be increased in steps of 5 dB or 10 dB until a maximum score has been found or the subject reports discomfort or fatigue. If the score decreases at higher levels (the roll-over effect), the test shall be continued at lower levels.

b) To determine a speech recognition score at the most comfortable loudness level, the test level shall be selected by presenting continuously a speech signal of the same type that is to be used for the actual test. Instruct the subject to report on its loudness (e.g. much too low, comfortable, too high, much too high). Often, halfway between too low and too high is a good initial choice for the most comfortable level.

c) To determine a half-optimum speech level, the maximum speech recognition score first has to be determined. Reduce the level in steps of 5 dB or 10 dB and present a full test list on each level until one level yields a score above half of the maximum score and the next lower level yields a score below half of the maximum score. Determine the integer value of the half-optimum speech level by means of linear interpolation between the latter two test levels.

Express the score obtained as a percentage and record the level at which it was achieved.

14 Contralateral masking

To avoid hearing the speech signals in the non-test ear during monaural speech audiometry, it may be necessary to apply air-conducted masking noise to the non-test ear. IEC 60645-2 specifies the reference conditions for one type of masking noise. The effective masking level for speech by a masking sound having a certain frequency spectrum and sound pressure level depends on the characteristics of the speech signal. Therefore, generally valid effective masking levels cannot be specified.

Masking shall be used in air-conduction testing if the speech signal level, expressed as the hearing level for speech, exceeds by 40 dB or more the average bone-conduction hearing threshold levels for two of the frequencies out of 500 Hz, 1 000 Hz, and 2 000 Hz of the contralateral ear showing the lowest hearing threshold levels.

The minimum masking level for air-conduction testing L_M, expressed as the effective masking level, shall be:

$$L_M = L_t - 40\ \text{dB} + (L_{Amin} - L_{Bmin})$$

where

L_t is the speech level, expressed as the hearing level for speech;

L_{Amin} is the average pure-tone air-conduction threshold level of the ear to be masked at two frequencies out of 500 Hz, 1 000 Hz, and 2 000 Hz, having the lowest hearing threshold levels, expressed as the hearing level;

L_{Bmin} is the average pure-tone bone-conduction threshold level of the ear to be masked at two frequencies out of 500 Hz, 1 000 Hz, and 2 000 Hz having the lowest hearing threshold levels, expressed as the hearing level.

In some cases, the masking level may have to be increased to ensure sufficient masking of the non-test ear.

15 Speech audiometry with competing sound

15.1 Type of competing sound

A frequency-weighted unmodulated random noise in accordance with IEC 60645-2 may be used as a competing sound. If an amplitude-modulated or any other noise or competing sound is used, its characteristics shall be specified. When other competing sounds than those included in the reference recording are used no valid reference values exist. See Annex B.

NOTE It is recognized that competing sounds using recordings of human speech can also be used, such as multitalker babble (cocktail-party noise) or connected speech from a single speaker. With such types of noise, larger variation in test results can occur than with weighted random noise.

15.2 Presentation of competing sound

In earphone testing, the competing sound is presented by means of the same earphone as used for the speech test signal. In sound field audiometry, the position of all loudspeakers in use shall be specified. The recommended position of the speech loudspeaker is frontal to the listener at an angle of incidence of 0°. Speech recognition with competing sound in a sound field may be affected by the acoustical characteristics of the test room and may therefore require reference data for each particular test environment.

It is recommended that the competing sound be presented by two loudspeakers symmetrically located at the angles of incidence of ±45°. The competing sound, if presented from two or more loudspeakers, shall be non-coherent.

15.3 Speech and competing sound levels

The levels of speech and competing sound shall be measured as defined in 3.17 and 3.30. The recommended speech level is 65 dB, which approximately corresponds to normal speech level in conversation. If other speech levels are used, they shall be clearly stated. The level of the competing sound may be either fixed or variable. If fixed, a noise level of 60 dB is recommended. If varied, it shall be changed in steps of 5 dB or less.

15.4 Test procedure

15.4.1 General

The test may be performed either to determine a speech recognition score at one or more fixed speech-to-noise ratios or to determine the speech-to-noise ratio at the speech recognition threshold.

15.4.2 Speech recognition score at fixed speech-to-noise ratio

15.4.2.1 Set the audiometric equipment to the required speech level.

15.4.2.2 Familiarize the test subject with the task by setting the competing sound level low (e.g. at 20 dB below the speech level) and presenting a number of test items which are clearly audible.

15.4.2.3 Set the competing sound level to the value required for the test and present a complete test list. Calculate the score as a percentage.

15.4.3 Speech recognition threshold with a competing sound

15.4.3.1 General

15.4.3.1.1 In the procedures described, increasing the level of the competing sound may be replaced by decreasing the level of the speech signal and vice versa. The test result may be influenced by the alternative chosen.

Alternative procedures are described for determining the speech-to-noise ratio that yield the speech recognition threshold at a given speech or competing sound level. They are expected to yield comparable results. However, experimental evidence for this is still unavailable. The procedures are identical from 15.4.3.1.2 to 15.4.3.1.3. Then, the procedure to decrease speech-to-noise ratio using 5 dB steps continues in 15.4.3.2 and the alternative procedure continues in 15.4.3.3.

15.4.3.1.2 Follow steps 15.4.2.1 and 15.4.2.2.

15.4.3.1.3 Increase the competing sound level in steps of 5 dB or less, and present at least two test items at each level until the test subject incorrectly recognizes one test item.

15.4.3.2 Procedure to decrease speech-to-noise ratio using 5 dB steps

15.4.3.2.1 Present a set of test items at this speech-to-noise ratio and record the number of correctly recognized test items. A set shall contain at least 10 test items.

15.4.3.2.2 If the subject scores at least 50 % on the set of test items, increase the competing sound level in steps of 5 dB or less and present a new set of test items on each speech-to-noise ratio until the subject scores less than 50 % on a set of test items. Determine the speech-to-noise ratio at the speech recognition threshold by means of linear interpolation between the lowest level that yielded more than 50 % correct and the highest level that yielded less than 50 % correct. If a score of exactly 50 % is obtained at one speech-to-noise ratio, that ratio represents the speech recognition threshold.

15.4.3.3 Alternative procedure to decrease speech-to-noise ratio

15.4.3.3.1 When one test item is missed, present a second with the same competing sound level. Continue this process in ascending steps of 10 dB until a level of the competing sound is reached at which two consecutive test items are missed at the same level.

15.4.3.3.2 Reduce the competing sound level by 10 dB. This defines the starting level of the competing sound, $L_{N,S}$.

15.4.3.3.3 Present two test items at this starting level and at each successive 2 dB increment if a step size of 2 dB is to be used. For a step size of 5 dB, see step 15.4.3.3.6.

15.4.3.3.4 Continue this process if at least five out of the first six test items are repeated correctly. If this criterion is not met, reduce the starting level by 4 dB to 10 dB.

15.4.3.3.5 The series of descending speech-to-noise ratio (ascending competing sound level) is terminated when the subject responds incorrectly to five of the last six test items presented.

15.4.3.3.6 If a step size of 5 dB is used, five test items should be presented at each level and the test terminated when all test items at a single level are incorrectly recognized.

15.4.3.3.7 For the selected speech level, calculate the sound pressure level of the competing sound at which 50 % of the test items can be correctly recognized according to the following equation:

$$L_{\mathrm{N,SRT}} = L_{\mathrm{N,S}} + \frac{dr}{n} - \frac{d}{2}$$

where

$L_{\mathrm{N,SRT}}$ is the competing sound level, in decibels, for 50 % correct speech recognition score;

$L_{\mathrm{N,S}}$ is the starting level of the competing sound, in decibels, according to step 15.4.3.3.2;

d is the step size, in decibels;

r is the total number of correct responses given from the starting level according to step 15.4.3.3.3 up to the termination according to step 15.4.3.3.5 or 15.4.3.3.6;

n is the number of test items per step.

Express the result as the difference between the speech level used and $L_{\mathrm{N,SRT}}$.

16 Format of speech audiogram

When a graphical representation of the results of speech audiometry (a speech audiogram) is to be used, it shall show the speech recognition score as a percentage along the ordinate, and the speech level, in decibels, or hearing level for speech, in decibels, or speech-to-noise ratio, in decibels, along the abscissa. The scale ratio should be 20 % corresponding to 10 dB. The type of speech material and any competing noise used shall be reported.

A speech audiogram should contain the corresponding reference speech recognition curve for the actual speech material used.

17 Measurement uncertainty

The uncertainty of speech recognition scores determined in accordance with any of the procedures specified in this part of ISO 8253 depends on a variety of parameters, such as:

a) the performance of the audiometric equipment used;

b) the type of transducers used and their fitting by the tester;

c) the speech material used;

d) the test procedure used;

e) the conditions of the test environment, especially the ambient noise;

f) the qualification and experience of the tester;

g) the co-operation of the test subject and the reliability of responses;

h) the use of non-optimized masking noise.

Due to the complexity of the measurement process including personal behaviour of both the test subject and the tester, it is nearly impossible to express the measurement uncertainty in a single generally valid figure.

However, a detailed evaluation of measurement uncertainty provides useful information on the reliability of audiometric test results and yields a sufficient estimate of the uncertainty in most applications.

The uncertainty of results of measurements according to this part of ISO 8253 shall be evaluated in compliance with ISO/IEC Guide 98-3. If reported, the expanded uncertainty together with the corresponding coverage

factor for a coverage probability of 95 %, as defined in ISO/IEC Guide 98-3, shall be given. Guidance on the determination of the expanded uncertainty is given in Annex E.

18 Maintenance and calibration of equipment

18.1 General

Correct calibration of audiometric equipment is highly important for reliable results. In order to ensure this, the following scheme, consisting of three stages of checks and calibration procedures, is recommended:

— stage A: routine checking and listening tests;

— stage B: periodic electroacoustic tests;

— stage C: basic calibration tests.

18.2 Intervals between tests

The recommended intervals at which the various tests are to be carried out are, by necessity, only a guide. They should be adhered to unless or until there is evidence that a different interval would be appropriate.

It is recommended that stage A tests be made each day before the equipment is to be used. It is recommended that tests of the sound pressure levels at the reference point in a sound field test facility should be made at intervals not exceeding 3 months and additionally whenever any alterations are made to the facility, e.g. after moving any furniture. The periodic electroacoustic tests, stage B, should be performed at intervals of 3 to 6 months, although different intervals may be acceptable in the light of experience with particular equipment and use, provided that the stage A tests are regularly and carefully applied. However, annual intervals should not be exceeded.

The basic calibration test, stage C, need not be used on a routine basis if stages A and B tests are regularly performed. A stage C test is required only when a serious equipment fault or error occurs or when, after a long period of time, it is suspected that the equipment may no longer be performing fully to specification. It is recommended that equipment should be submitted for a stage C test at intervals not exceeding 5 years.

18.3 Stage A: Routine checking and listening tests

18.3.1 The purpose of routine checking is to ensure, as far as possible, that the equipment is working properly and that its calibration has not noticeably altered. The ambient noise conditions during the test shall be comparable to those in force when the equipment is in normal use.

The test procedure is specified in 18.3.2 to 18.3.9.

18.3.2 Examine and clean, as appropriate, the equipment and all accessories. Check earphone cushions, plugs, and leads for signs of wear and damage. Damaged or badly worn parts should be replaced.

18.3.3 Switch on the equipment and leave it on for the recommended warm-up time or at least 5 min. Carry out any set-up procedures as specified by the manufacturer. Ensure that transducer serial numbers correspond to the instrument serial number, if possible.

18.3.4 If sound field testing is being undertaken, ensure that the reference point is at the correct position and is clearly identified.

18.3.5 Check that the audiometer output is approximately correct and that ambient noise levels are typical and acceptable by having a person listen to low-level speech test signals. This person should preferably be the same each time the test is performed and shall have well-known hearing threshold levels within the normal range. The test shall be performed with all appropriate output transducers.

18.3.6 Listen to the speech test material at a higher level (e.g. at a hearing level for speech of 60 dB to 70 dB) on all appropriate functions and with all test signals available. Listen for proper functioning, in particular for the absence of distortion and freedom from interfering noise.

18.3.7 Listen at low levels and ensure that no hum or noise or any other unwanted sound from the equipment is audible at the test subject's position.

18.3.8 Check that the attenuators do attenuate the signals over their full range.

18.3.9 Ensure that the subject's talk-back system and the monitor circuits operate correctly.

18.4 Stage B: Periodic electroacoustic tests

Periodic electroacoustic tests consist of measuring and comparing results of the following:

a) the frequency response for the equipment including all appropriate output transducers as specified in IEC 60645-2;

b) the output levels of the transducers as specified in IEC 60645-2;

c) the attenuator steps (over a significant part of the range) in accordance with IEC 60645-1;

d) the harmonic distortion in accordance with IEC 60645-2;

e) the masking noise levels in accordance with IEC 60645-2;

f) the headband force of the transducers.

In addition, a routine stage A check shall be performed.

18.5 Stage C: Basic calibration tests

A basic calibration shall ensure that the audiometric equipment, the sound field, if appropriate, and the ambient noise levels meet all relevant specifications in IEC 60645-2 and this part of ISO 8253.

Annex A
(informative)

Example of speech materials

A.1 General

Speech materials can be divided into different classes. This annex lists some typical examples of speech materials for these different classes. Because of the large number of speech materials, it is impossible to give a complete list.

A.2 Examples of open-set materials

Examples of open-set materials are given in References [12]–[15] and [30].

A.3 Examples of closed-set materials

Examples of closed-set materials are given in References [16]–[18] and [29].

A.4 Examples of open-set sentence tests

Examples of open-set sentence materials are given in References [19]–[24] and [28].

A.5 Examples of closed-set sentence tests

Examples of closed-set sentence materials are given in References [25]–[27].

Annex B
(informative)

Examples of competing sound conditions

B.1 General

If competing sound(s) are included in the reference recording, for which speech recognition reference values are available, those sounds are preferred.

B.2 Non-modulated noises

This kind of competitive sound shows no or only a very small amount of amplitude modulations. In many cases, these noises represent the mean long-term spectrum of the speech material of the speech recognition test. Usually, this kind of noise yields a high test-retest reliability of the speech recognition threshold level. However, the differences between speech recognition threshold levels for different degrees of hearing loss are relatively small. Typical examples are the CCITT noise (ITU-T Recommendation G.227[8]; telephone frequency characteristic) or the ICRA1 noise (Reference [31]) which represents the mean speech spectrum averaged across different languages.

B.3 Modulated noises

These sounds show amplitude modulations similar to single talkers or groups of several talkers. One example is the ICRA5 noise (Reference [32]) which represents the mean long-term spectrum of speech and the time structure of one interfering speaker. The amplitude modulations are independent in three different frequency channels. The modified version ICRA5-250 (Reference [32]) shows a modified time structure as the pauses were limited to 250 ms. Usually, this kind of noise yields larger differences between speech recognition threshold levels for different degrees of hearing loss compared to non-modulated noises. A similar noise is the noise described in Reference [33]. In this noise the amplitude modulations are the same for all frequencies.

B.4 Babble noises

This kind of competitive sound consists of one or more superposed speakers. These kinds of noises show different degrees of modulation and different degrees of informational masking.

All different types of interfering signals (reversed speech, running speech, babble noises composed of several speakers, etc.) may be applied in speech recognition tests in noise as well.

The presentation of the interfering sound can differ in the way of presentation. In the gated condition, the noise begins several milliseconds (e.g. 500 ms) before the beginning of the speech and ends several milliseconds after the end of the speech. In the continuous condition, the interfering sound is presented during the whole test procedure.

Since different interfering signals yield different speech recognition reference values and test-retest reliabilities the interfering noise should be reported in the speech audiogram (see Clause 16).

Annex C
(informative)

Typical results

C.1 Typical reference speech recognition curves

As an example, data for the Göttingen sentences (Reference [20]) are shown. The speech material was based on 20 lists of 10 sentences each, and tested on 12 normal-hearing subjects. The speech recognition score (between 0 % and 100 %), R, is given by

$$R = \frac{100}{1 + \exp\left[0{,}04 \times S\left(L_{50} - L\right)\right]}$$

where

L is the speech level in quiet or SNR in competing noise;

L_{50} is the speech recognition threshold in dB (50 % correct).

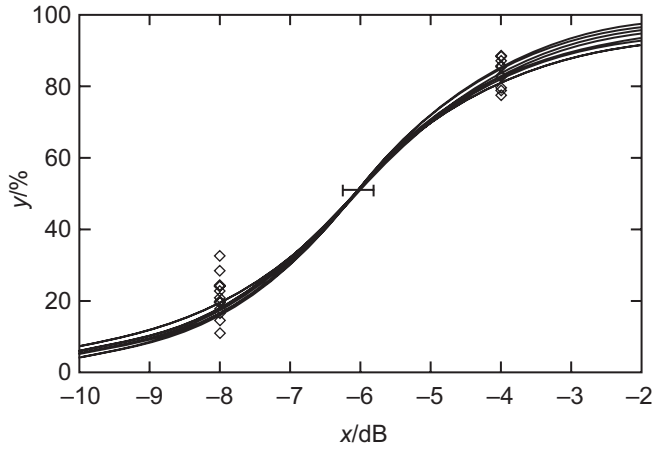
NOTE For tests without competing noise: L_{50} = 20 dB. For tests with speech simulating unmodulated noise, the speech-to-noise ratio is $L_{50} = -6$ dB.

S is the slope of the recognition curve in %/dB at L_{50}, and

S_S = 11 %/dB in silence,

S_N = 19 %/dB in noise.

Figure C.1 shows a typical reference speech recognition curve using Göttingen sentences in competing noise.



Key

y proportion of correct words

x speech-to-noise ratio

Figure C.1 — Typical reference speech recognition curve using Göttingen sentences in competing noise

NOTE The symbols at −4 and −8 dB SNR denote the intelligibility scores for each of the 20 lists averaged across subjects. The error bar at −6 dB denotes the standard deviation of the list-specific speech recognition threshold.

C.2 Typical speech recognition threshold levels

Table C.1 displays typical speech recognition threshold levels yielded using some of the speech recognition tests given in Annex A for otologically normal listeners.

Table C.1 — Typical speech recognition threshold levels

Type of test material	Reference	Scoring	Speech recognition threshold level for otologically normal listeners
Open-set sentences in silence	[20]	Words	20 dB
Open-set sentences in noise	[20]	Words	SNR: −6 dB
Open-set sentences in noise	[19]	Words	SNR: −5 dB
Open-set sentences in noise	[19]	Whole sentences	SNR: −3 dB
Open-set sentences in noise	[23]	Whole sentences	SNR: −6 dB
Open-set sentences in noise	[22]	Whole sentences	SNR: −3 dB
Open-set words in silence	[13]	Words	28 dB
Open-set words in silence	[30]	Words	23 dB
Open-set words in silence	[13]	Phonemes	24 dB
Closed-set sentences in noise	[27]	Words	SNR: −8 dB

NOTE All tests of speech recognition in noise in these examples used spectrally matched random noise as competing sound.

C.3 Typical test-retest reliabilities

Table C.2 displays typical test-retest reliabilities for speech recognition scores yielded using some examples of the speech materials given in Annex A.

Table C.2 — Typical test-retest reliabilities

Type of test material, number of test items	Reference	Scoring	Typical test-retest reliabilities for speech recognition threshold level and speech recognition score, respectively, for otologically normal listeners
Open-set sentences in silence, 10 sentences	[20]	Words	2 dB
Open-set sentences in noise, 10 sentences	[19]	Whole sentences	0,9 dB
Open-set sentences in noise, 10 sentences	[21]	Whole sentences	1,2 dB
Open-set words in silence, 50 words	[13]	Words	10 %
Open-set words in silence, 50 words	[13]	Phonemes	8 %
Closed-set sentences in noise, 20 sentences	[27]	Words	1 dB

Annex D
(informative)

Optimization of perceptual balance of test lists

The perceptual equivalence of test lists may be optimized in different ways.

One possibility is to adjust the level of the particular test items with respect to the measured average speech recognition scores of the items. Items with lower speech recognition compared to the average are amplified and items with higher speech recognition compared to the average are attenuated. The level adjustments are only applied for particular speech recognition threshold levels with differences exceeding 0,5 dB compared to the average speech recognition threshold level of the speech material. The level adjustments shall not yield perceivable loudness differences between the particular test items within the test stimuli (e.g. obvious loudness differences between the particular words of a test sentence).

Another possibility is to leave the levels of the different test items within one speech presentation (e.g. sentence) unchanged and to introduce weighting factors for the particular test items. Items with lower speech recognition compared to the average get higher weights and items with higher speech recognition compared to the average get lower weights. The sum of weighting factors per test stimulus (e.g. a sentence) shall be 1.

Annex E
(informative)

Measurement uncertainty

E.1 General

The generally accepted format for the expression of uncertainties associated with results of measurements is that given in ISO/IEC Guide 98-3. This format requests a functional relationship (model function) to be established between the measurand, which in the context of this part of ISO 8253 is the speech recognition threshold level of a test subject, and several input quantities describing effects that may influence the measurement result. Each of these input quantities is characterized by its estimate, its probability distribution and its standard uncertainty. The existing knowledge on these input quantities is to be compiled in an uncertainty budget from which the combined standard uncertainty and the expanded uncertainty of the measurement result can be derived.

Scientifically verified data necessary to establish a sound uncertainty budget for each measurement performed using any of the procedures specified in this part of ISO 8253 are, at the time of publication, not available. However, an indication of the relevant sources of uncertainty and their characteristics can be given, mostly based on empirical knowledge. The general approach to the calculation of uncertainties conforming with ISO/IEC Guide 98-3 is illustrated. It allows an approximate determination of uncertainties under special assumptions.

E.2 Model function

The expression for the determination of the speech recognition threshold level L_{HT} is given by Equation (E.1):

$$L_{HT} = L'_{HT} + \delta_{eq} + \delta_{tr} + \delta_{m} + \delta_{te} + \delta_{su} + \delta_{pr} + \delta_{tm} + \delta_{n} + \delta_{mth} \quad \text{(E.1)}$$

where

L'_{HT} is the outcome of a speech recognition threshold level determination in accordance with any of the procedures specified in this part of ISO 8253 (for details see E.3.2);

δ_{eq} is an input quantity to allow for any deviation from nominal performance of the audiometric equipment used (E.3.3);

δ_{tr} is an input quantity to allow for uncertainties due to the use of a certain type of transducer and its fitting (E.3.4),

δ_{m} is an input quantity to allow for any uncertainties due to a non-optimized masking or competing noise (E.3.5);

δ_{te} is an input quantity to allow for any uncertainties due to lacking qualification and experience of the tester (E.3.6);

δ_{su} is an input quantity to allow for any uncertainties due to lacking co-operation and non-reliable responses of the test subject (E.3.7);

δ_{pr} is an input quantity to allow for any uncertainties due to special problems arising from an unusually difficult measurement situation (E.3.8);

δ_{tm} is an input quantity to allow for different performance of the speech material recording according to Clause 4 (E.3.9);

δ_n is an input quantity to allow for the influence of non-ideal environmental conditions, especially ambient noise (E.3.10);

δ_{mth} is an input quantity to allow for any uncertainties due to the test method used (E.3.11).

Usually, each of the δ input quantities is considered to have an estimate of 0 dB, i.e. no correction is applied to the determined hearing threshold level. However, each of these quantities is associated with an uncertainty as explained in E.3. None of the input quantities is correlated to any other to any significant extent.

NOTE It is possible that, for other measurands, Equation (E.1) requires appropriate modification.

E.3 Input quantities

E.3.1 General

Input quantities described in E.3.2 to E.3.6 and E.3.9 to E.3.11 have to be considered in nearly all audiometric applications while those described in E.3.7 to E.3.8 are to be taken into account only in exceptional situations subject to the personal judgement of the tester.

E.3.2 Determined hearing threshold level, L'_{HT}

During routine speech audiometry, the speech recognition threshold level of a test subject is usually determined just once for each ear or once in a binaural test. However, based on empirical knowledge, the following approximate standard uncertainties for repeated measurements after proper training and under identical test conditions can be assumed (see 5.3 and C.3):

— for air conduction testing using one list of 10 sentences: 1 dB;

— for air conduction testing using a set of 10 single test words: 2,5 dB.

The probability distribution of probable values of L'_{HT} can be assumed to be normal, its estimate is designated $L'_{HT,est}$ (see Table E.1).

E.3.3 Audiometric equipment, δ_{eq}

Assuming that the audiometric equipment meets the requirements of IEC 60645-2 for a type A or B audiometer, its dominant contribution to the measurement uncertainty is probably given by the deviation of the output levels provided from nominal values. IEC 60645-2 specifies the following maximum deviations:

— air conduction: ±3 dB.

Unless any more specific information on the performance of the equipment is available, the probability distribution of the output levels can be assumed to be rectangular, resulting in standard uncertainties that are equal to half the maximum spread of possible values, divided by √3.

If the step size of the hearing level control is 5 dB, this introduces another non-negligible uncertainty contribution with a rectangular probability distribution and a standard uncertainty of 2,5/√3 dB.

The two contributions result in an approximate overall standard uncertainty, eg. for air conduction, of

$$\sqrt{\left(3/\sqrt{3}\right)^2+\left(2{,}5/\sqrt{3}\right)^2}\ \text{dB} = 2{,}3\ \text{dB}\ .$$

E.3.4 Transducers and their fitting, δ_{tr}

The sound pressure levels provided by different types of transducers to the ear of a test subject may be differently sensitive to anatomic and physiological characteristics of the test subject, to their placement at the ear and to deviations of the headband forces from nominal values. Generally valid figures on the uncertainty

contribution from these effects cannot be stated at present. However, unless more detailed knowledge is available, a standard uncertainty of 2,5 dB may be assumed. The two effects together result in an approximate standard uncertainty of $\sqrt{(1{,}5^2 + 2{,}5^2)}$ dB = 2,9 dB.

E.3.5 Masking and competing noise, δ_m

Measured speech recognition threshold levels may be affected by the use of non-optimized masking noise. No generally valid figures on the contribution of this effect to the measurement uncertainty can be given. However, a standard uncertainty of 2 dB may provisionally be attributed to δ_m with a normal probability distribution if masking noise is applied. The effect of variations in a competing noise, assuming the same type of noise and other test conditions, is included in the standard uncertainty for repeated measurements.

E.3.6 Experience of the tester, δ_{te}

For a qualified tester with sufficient experience, the contribution to uncertainty due to personal judgements may be considered to be included in the standard uncertainty for repeated measurements (E.3.2) for usual test situations. Under special circumstances, it may however be appropriate to attribute an additional uncertainty to δ_{te}.

E.3.7 Responses of the test subject, δ_{su}

Under usual circumstances, uncertainties due to minor inconsistencies in the responses of a test subject are included in the standard uncertainty for repeated measurements (E.3.2). There may, however, be reasons to attribute an additional standard uncertainty to δ_{su} in exceptional situations.

E.3.8 Special measurement situations, δ_{pr}

There may be exceptional cases where it is extremely difficult to determine the speech recognition threshold level of a test subject. In such cases, an additional standard uncertainty may be attributed to δ_{pr}.

E.3.9 Different performance of the test material, δ_{tm}

Measured speech recognition threshold levels may be affected by different performance of the test material recording according to Clause 4. No generally valid figures on the contribution of this effect to the measurement uncertainty can be given. However, a standard uncertainty of 2 dB may provisionally be attributed to δ_{tm} with a normal probability distribution. The effect of variations in performance of the test material recording are included in the standard uncertainty for repeated measurements when using different test lists of the same speech material.

E.3.10 Environmental conditions, δ_n

If the requirements on ambient noise (see Clause 7 are fully met, the standard uncertainty of δ_n may be assumed to be 2 dB with a normal probability distribution when determining speech detection threshold, considering test subjects with a hearing threshold level close to 0 dB. For determination of speech recognition scores and for test subjects with hearing threshold levels significantly above 0 dB, the uncertainty contribution due to ambient noise may be negligible.

E.3.11 Test method used, δ_{mth}

Measured speech recognition threshold levels are affected by the test method used as specified in 12.1 to 12.5. Besides others, sources of uncertainty are the method itself (descending versus adaptive), the attenuator step size used, and the number of test items used (complete lists versus sets of test items or single test items). As stated in Clause 12, no experimental evidence on the equivalence of the various methods exists at the time of publication. Therefore, only a very rough estimate of the standard uncertainty to be expected may be given. However, 2 dB might be a realistic value for provisional use.

E.4 Uncertainty budget

The contributions to the combined uncertainty associated with the value of the determined hearing threshold level depend on the standard uncertainties, u_i, as described in E.3 and the related sensitivity coefficients, c_i. The sensitivity coefficients are a measure of how the values of the hearing threshold level are affected by changes in the values of the respective input quantities. Mathematically they are equal to the partial derivative of the model function with respect to the relevant input quantity. The contributions of the respective input quantities are then given by the products of the standard uncertainties and their associated sensitivity coefficients. The uncertainty budget compiles the available information on the various uncertainty contributions in tabular form.

Table E.1 — General form of an uncertainty budget for hearing threshold level determinations

Quantity	Estimate dB	Standard uncertainty u_i dB	Probability distribution	Sensitivity coefficient c_i	Uncertainty contribution $c_i u_i$ dB
L'_{HT}	$L'_{\mathrm{HT,est}}$	u_1	normal	1	u_1
δ_{eq}	0	u_2	rectangular	1	u_2
δ_{tr}	0	u_3	normal	1	u_3
δ_{m}	0	u_4	normal	1	u_4
δ_{te}	0	u_5	normal	1	u_5
δ_{su}	0	u_6	normal	1	u_6
δ_{pr}	0	u_7	normal	1	u_7
δ_{tm}	0	u_8	normal	1	u_8
δ_{n}	0	u_9	normal	1	u_9
δ_{mth}	0	u_{10}	normal	1	u_{10}

E.5 Combined and expanded uncertainty

The combined uncertainty for the speech recognition threshold level is given by Equation (E.2):

$$u = \sqrt{\sum_{i=1}^{10} u_i^2} \tag{E.2}$$

ISO/IEC Guide 98-3 requires an expanded uncertainty, U, to be specified, such that the interval $[L_{\mathrm{HT}} - U, L_{\mathrm{HT}} + U]$ covers for example 95 % of the values of L_{HT} that might reasonably be attributed to L_{HT}. For this purpose, a coverage factor k is used such that $U = k\,u$. For a coverage probability of 95 % and for a normal distribution k has a value of 2.

E.6 Example

As an example, the expanded measurement uncertainty is evaluated for the determination of the speech recognition threshold level of a test subject using air conduction audiometry without masking and assuming that the requirements on ambient noise are met and no further uncertainty contribution arise from any other sources. The uncertainty budget then has the form listed in Table E.2.

Table E.2 — Example of an uncertainty budget for the measurement conditions stated above

Quantity	Estimate dB	Standard uncertainty dB	Probability distribution	Sensitivity coefficient	Uncertainty contribution dB
L'_{HT}	$L'_{HT,est}$	2,5	normal	1	2,5
δ_{eq}	0	2,3	rectangular	1	2,3
δ_{tr}	0	2,9	normal	1	2,9
δ_{mth}	0	2	normal	1	2

Combined standard uncertainty: $u = 4{,}9$ dB.

Expanded measurement uncertainty for 95 % coverage probability, rounded to the nearest full decibel: $U = 10$ dB.

Bibliography

[1] ISO 389-1, *Acoustics — Reference zero for the calibration of audiometric equipment — Part 1: Reference equivalent threshold sound pressure levels for pure tones and supra-aural earphones*

[2] ISO 389-4, *Acoustics — Reference zero for the calibration of audiometric equipment — Part 4: Reference levels for narrow-band masking noise*

[3] ISO/TR 25417, *Acoustics — Definitions of basic quantities and terms*

[4] IEC 50801, *International Electrotechnical Vocabulary — Chapter 801: Acoustics and electroacoustics*

[5] IEC 60318-1, *Electroacoustics — Simulators of human head and ear — Part 1: Ear simulator for the measurement of supra-aural and circumaural earphones*

[6] IEC 60318-3, *Electroacoustics — Simulators of human head and ear — Part 3: Acoustic coupler for the calibration of supra-aural earphones used in audiometry*

[7] IEC 61260, *Electroacoustics — Octave-band and fractional-octave-band filters*

[8] ITU-T Recommendation G.227, *International analogue carrier systems — General characteristics common to all analogue carrier-transmission systems — Conventional telephone signal*. Available (viewed 2012-02-21) at: http://www.itu.int/rec/T-REC-G.227-198811-I/en

[9] LUDVIGSEN, C. Comparison of certain measures of speech and noise level. *Scand. Audiol*. 1992, **21**, pp. 23–29

[10] MARTIN, M., editor. *Speech audiometry*. London: Whurr, 1987

[11] WILSON, R.H., MORGAN. D.E., DIRKS, D.D. A proposed SRT procedure and its statistical precedent. *J. Speech Hear. Dis.* 1973, **38**, pp. 184–191

[12] ELBERLING, C., LUDVIGSEN, C., LYREGAARD, P.E. Dantale: a new Danish speech material. *Scand. Audiol*. 1989, **18**, pp. 169–175

[13] KEIDSER, G. Normative data in quiet and in noise for “DANTALE” — a Danish speech material. *Scand. Audiol*. 1993, **22**, pp. 231–236

[14] OLSEN, S.Ø. Evaluation of the list of numerals in the Danish speech audiometry material. *Scand. Audiol.* 1996, **25**, pp. 103–107

[15] SAKAMOTO, S., YOSHIKAWA, T., AMANO, S., SUZUKI, Y., KONDO, T. New 20-word lists for word intelligibility test in Japanese, In: *Proc. of 9th International Conference on Spoken Language Processing (INTERSPEECH 2006 - ICSLP)*, 2006, pp. 2158–2161

[16] BRAND, T., ACHTZEHN, J., KOLLMEIER, B. Erstellung von Testlisten für den Oldenburger Kinder-Reimtest [Development of test lists for the Oldenburger children’s rhyme test]. *Z. Audiol.* 1999, (Suppl. II), pp. 50–51

[17] BRAND, T., WAGENER, K.C. Wie lässt sich die maximale Verständlichkeit optimal bestimmen? [How can maximum speech recognition be determined optimally?] In: *8. Jahrestagung der Deutschen Gesellschaft für Audiologie*, Tagungs CD, ISBN 3-9809869-4-2. 2005

[18] VON WALLENBERG, E.L., KOLLMEIER, B. Sprachverständlichkeitsmessungen für die Audiologie mit einem Reimtest in deutscher Sprache: Erstellung und Evaluation von Testlisten [Measurements of speech recognition in audiology with a rhyme test in the German language: Development and evaluation of test lists]. *Audiol. Akust.* 1989, **28**, pp. 50–65

[19] HÄLLGREN, M., LARSBY, B., ARLINGER, S. A Swedish version of the Hearing in Noise Test (HINT) for measurement of speech recognition. *Int. J. Audiol*. 2006, **45**, pp. 227–237

[20] KOLLMEIER, B., WESSELKAMP, M. Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment. *J. Acoust. Soc. Am*. 1997, **102,** pp. 2412–2421

[21] KALIKOW, D.N., STEVENS, K.N., ELLIOT, L.L. Development of a test of speech intelligibility in noise using sentences with controlled word predictability. *J. Acoust. Soc. Am*. 1977, **61**, pp. 1337–1351

[22] NILSSON, M., SOLI, S.D., SULLIVAN, J.A. Development of Hearing In Noise Test for measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. Am*. 1994, **95**, pp. 1085–1099

[23] PLOMP, R., MIMPEN, A.M. Improving the reliability of testing the speech reception threshold for sentences. *Audiology* 1979, **18**, pp. 43–52

[24] KILLION, M.C., NIQUETTE, P.A., GUDMUNDSEN, G.I., REVIT, L.J., BANERJEE, S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *J. Acoust. Soc.Am*. 2004, **116**, pp. 2395–2405

[25] HAGERMAN, B. Sentences for testing speech intelligibility in noise. *Scand. Audiol*. 1982, **11**, pp. 79–87

[26] WAGENER, K., BRAND, T., KÜHNEL, V., KOLLMEIER, B. Entwicklung und Evaluation eines Satztests für die deutsche Sprache I-III: Design, Optimierung und Evaluation des Oldenburger Satztests [Development and evaluation of a sentence test for the German language I-III: Design, optimization and evaluation of the Oldenburger sentence test]. *Z. Audiol*. 1999, **38**, pp. 4–15, 44–56, 86–95

[27] WAGENER, K., JOSVASSEN, J.L., ARDENKJAER, R. Design, optimization and evaluation of a Danish sentence test in noise. *Int. J. Audiol*. 2003, **42**, pp. 10–17

[28] WAGENER, K.C., KOLLMEIER, B. Evaluation des Oldenburger Satztests mit Kindern und Oldenburger Kinder-Satztest [Evaluation of the Oldenburger sentence test with children and the Oldenburger sentence test for children]. *Z. Audiol.* 2005, **44**, pp. 134–143

[29] WAGENER, K.C., BRAND, T., KOLLMEIER, B. Evaluation des Oldenburg Kinder-Reimtests in Ruhe und im Störgeräusch [Evaluation of the Oldenburger rhyme test for children in silence and in noise]. *HNO* 2006, **54**, pp. 171–178

[30] HAN, H., LEE, J., CHO, S., KIM, J., LEE, K., CHOI, W. Reference sound pressure level for Korean speech audiometry. *Int. J. Audiol*. 2011, **50**, pp. 59–62

[31] DRESCHLER, W.A., VERSCHUURE, H., LUDVIGSEN, C., WESTERMANN, S. ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium of Rehabilitative Audiology. *Audiology* 2001, **40**, pp. 148–157

[32] WAGENER, K.C., BRAND, T. Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters. *Int. J. Audiol*. 2005, **44**, pp. 144–157

[33] FASTL, H. A background noise for speech audiometry. *Audiol. Acoust.* 1987, **26**, pp. 2–13

ICS 13.140

Price based on 31 pages

Copyrighted material licensed to Dublin Institute of Technology by SAI Global (www.saiglobal.com), downloaded on 12 Jul 12 by Ann McSweeney.
No further reproduction or distribution is permitted. Uncontrolled when printed.