INTERNATIONAL STANDARD

ISO 8253-3

Third edition
2022-01

Acoustics — Audiometric test methods —

Part 3: Speech audiometry

Acoustique — Méthodes d'essais audiométriques —

Partie 3: Audiométrie vocale

Reference number
ISO 8253-3:2022(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: 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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 43, *Acoustics*, in collaboration with the European Committee for Standardization (CEN) Technical Committee CEN/TC 211, *Acoustics,* in accordance with the Agreement on technical cooperation between ISO and CEN (Vienna Agreement).

This third edition cancels and replaces the second edition (ISO 8253-3:2012), which has been technically revised. The main changes compared to the previous edition are as follows:

— The technical requirements for recording of speech material were adapted to recent equipment and technology.

— The determination of reference speech recognition curves was revised. An annex that gives advices on how to determine the minimum number of subjects was introduced.

— The determination of speech recognition threshold levels is described in a more general manner.

— Symbols for the graphical representation of speech audiometry results were introduced.

A list of all the parts in the ISO 8253 series can be found on the ISO website.

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 document sets conditions for speech materials in order to ensure 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 document specifies basic methods for speech recognition tests for audiological applications.

NOTE Examples of speech materials are given in Annex A.

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

Furthermore, this document also specifies the determination of reference values and requirements for the realization and manner of presentation. In addition, there are features of speech tests described which are important to be specified, but which are not understood as a requirement.

This document specifies procedures and requirements for speech audiometry with the recorded test material being presented by an audiometer through a transducer, e.g., an earphone, bone vibrator, or loudspeaker arrangement 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 modified test procedures not specified in this document.

Specialized tests, such as those used for evaluating directional hearing and dichotic hearing, are outside the scope of this document.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements 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 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 60645-1:2017, *Electroacoustics — Audiometric equipment — Part 1: Equipment for pure-tone and 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.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

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

Note 1 to entry: A speech signal can be a voice signal or an acoustic signal simulating a voice signal.

3.2
test item
particular monosyllabic or polysyllabic word, or *logatom* (3.12), or spondee, 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 1 to entry: Scoring may be based on a complete test item or parts thereof being correctly recognized.

3.3
test list
number of selected *test items* (3.2), presented and scored as a single unit

3.4
set of test items
selected number of *test items* (3.2) from a test list

3.5
speech material
entire *set of test items* (3.4) which is used for speech recognition tests

Note 1 to entry: Usually the speech material is subdivided into several test lists.

3.6
open-set test
test in which the number of alternative responses to each *test item* (3.2) is unlimited

3.7
closed-set test
test in which the number of alternative responses to each *test item* (3.2) is limited

3.8
phoneme
the smallest phonetic unit that distinguishes one word from another in a particular language

3.9
phoneme class
subdivision of phonemes that show characteristic similarities in vocal production mode as well as in acoustical signal properties

3.10
phoneme distribution
relative distribution of the various phonemes within a given *speech material* (3.5)

3.11
syllable
segment of speech consisting of an (optional) onset, a nucleus, and an (optional) coda, where the nucleus most often consists of a vowel sound and the onset and coda most often consist of one or more consonant sounds

3.12
logatom
syllabic unit that has no verbal meaning to the listener

Note 1 to entry: A logatom is sometimes called a "nonsense syllable".

3.13
carrier phrase
sentence or phrase in which a *test item* (3.2) is embedded such that the correct recognition of the *test item* (3.2) is not dependent upon the context or meaning of the sentence or phrase

3.14
connected speech
continuous speech with natural intonation, consisting of consecutive sentences with logical connections

3.15
reference recording
master recording
well-defined recording that represents the *speech material* (3.5) and that is used for validation and application of the *speech material* (3.5)

3.16
speech level
sound pressure level of a specified *speech material* (3.5) as measured in an appropriate acoustic coupler, ear simulator or in a sound field, with a specified frequency and time weighting

Note 1 to entry: The frequency and the time weighting should be according to IEC 61672-1.

3.17
speech detection threshold level
for a given test subject, for a specified *speech material* (3.5) and a specified manner of signal presentation, the speech level at which it is detected as sound (not necessarily understood) in 50 % of the trials

Note 1 to entry: Is synonymous with the expression "speech awareness threshold".

3.18
speech recognition score
for a given test subject, for a specified *speech material* (3.5), a specified manner of signal presentation and at a specified speech level, the percentage of correctly recognized *test items* (3.2) or scorable items if the scoring method is not based on whole *test items* (3.2)

3.19
speech audiogram
graphical representation of the results of speech audiometry, where the speech recognition score is plotted as percentage along the ordinate and speech level, in dB, or hearing level for speech, in dB, or speech-to-noise level difference, in dB, along the abscissa

3.20
maximum speech recognition score
for a given test subject, for a specified *speech material* (3.5) and a specified manner of signal presentation, the maximum speech recognition score obtained regardless of the speech level

3.21
speech recognition threshold level
for a given test subject, for a specified *speech material* (3.5) and a specified manner of signal presentation, the lowest speech level at which the speech recognition score is equal to 50 % or another specified value

Note 1 to entry: Speech recognition threshold has been called "speech reception threshold".

3.22
reference speech recognition threshold level
for a specified *speech material* (3.5) 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 *speech material* (3.5) is appropriate

3.23
hearing level for speech
for a specified *speech material* (3.5) and a specified manner of signal presentation, the speech level minus the appropriate reference speech recognition threshold level

3.24
optimum speech level
for a given test subject, for a 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 a 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
most comfortable level for speech
for a given test subject, a specific *speech material* (3.5) and a specified manner of presentation, the speech level at which the loudness of the *speech signal* (3.1) is judged by the test subject to be most comfortable

3.27
speech recognition curve
for a given test subject, for a specified *speech material* (3.5) and a specified manner of signal presentation, a curve that describes the speech recognition score as a function of speech level

3.28
reference speech recognition curve
for a specified *speech material* (3.5) and a specified manner of presentation, a curve that describes the 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 *speech material* (3.5) is appropriate

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

3.30
competing sound level
sound pressure level of a competing sound as measured in an appropriate acoustic coupler, ear simulator or in a sound field, with a specified frequency and time weighting

Note 1 to entry: The frequency and the time weighting should be according to IEC 61672-1.

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

3.32
effective masking level for speech
level of a specified masking sound equivalent to that hearing level for speech to which the speech recognition threshold level for a specified *speech signal* (3.1) would be raised by the presence of that masking sound

4 Requirements for recording of speech material

4.1 General requirements

Each recording shall contain the following signals:

a) the speech material;

b) a signal for the calibration of the equipment for speech audiometry.

4.2 Calibration signal

The calibration signal shall meet the following requirements:

a) A minimum duration of 60 s.

b) The calibration signal for sound-field speech audiometry shall be a weighted random noise, e.g., as specified in IEC 60645-1:2017, 6.5.3, 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) The calibration signal for earphones or bone vibrators shall be as specified in b). Alternatively, a pure-tone signal may be used.

4.3 Reference recording

The voice characteristics of the speaker influence the test results. For general-use tests, a suitable speaker, either male or female, shall have a normal and clear articulation and should speak with a generally well understood dialect. The speaker shall be instructed to maintain clarity, natural pace, and vocal effort and avoid emphasis on keywords.

NOTE 1 Tests with representative speech facilitate the comparison of results.

NOTE 2 For some applications, e.g., localized speech tests, a specific dialect can be required.

A reference recording mainly consists of the elements specified in 4.1. The reference recording shall not be modified. All speech materials used for testing shall be equivalent copies of 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 sound or sounds, either on the same or on separate channels.

4.4 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 should have a reverberation time of less than 0,5 s at any frequency in the range from 125 Hz to 8 000 Hz.

4.5 Recording setup

The frequency response of the recording setup, including the microphone used, shall be documented along with the recording setup at least 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.

The signal-to-noise ratio of the recording shall be at least 40 dB when measured using the same frequency weighting as for measuring the speech level.

4.6 Signals and levels

A speech test may comprise different types of signals. Common signals are calibration signals, speech signals, announcing signals and competing sound signals.

NOTE 1 Not all speech tests comprise all of these signal types. The minimum type expected to be comprised is speech signals.

NOTE 2 Speech signals can be logatomes, syllables, words, or sentences.

4.6.1 Calibration signals

Calibration signals may comprise a signal used to determine the sound pressure level (referred to as the calibration signal) and additional signals used to check, for instance, the frequency response of the equipment. A speech test should not comprise more than one calibration signal to determine the sound pressure level.

NOTE 1 A common calibration procedure is to play the calibration signal in a defined state of the audiometry software and to adjust the equipment until a particular sound pressure level is reported by a suitable level meter. After that, the software is able to set the output to any required sound pressure level by adjusting either the signal amplitude or the audiometer level (attenuator or amplifier level).

If a speech test comprises a calibration signal, the calibration signal level shall be stated and measured as equivalent continuous sound pressure level (L_{eq}) with one of the frequency weightings specified in IEC 61672-1. For digital calibration signals, its root mean square (RMS) level re. full scale shall be documented within the speech test.

NOTE 2 Due to common speech spectra shapes, frequency weighting Z (linear) is expected to yield the same levels as weighting C. However, weighting C can be preferred in order to avoid influences of unheard interfering noise (with very low or very high frequencies).

NOTE 3 Not all sound level meters allow to determine L_{eq}, but instead provide particular time weightings (e.g., exponential "FAST" or "SLOW"). For an unmodulated or little modulated calibration signal, a slow time weighting can be a good approximation of L_{eq}.

If a speech test does not comprise a calibration signal, a particular part (subset) of the speech signals shall be specified which is to be used to determine the speech level in a calibration procedure. In this case, the speech level shall be stated and measured in the same way as defined above for the calibration signal level.

NOTE 4 When speech signals are used for calibration, a time weighting of the sound level meter is not a good approximation of L_{eq}.

4.6.2 Speech signals

The relation (amplitude ratio or level difference in dB, respectively) of all speech signals to the signal used for the calibration shall be documented within the speech test. The nominal speech level is given by the level of the calibration signal when played at a nominal level difference of 0 dB.

EXAMPLE If a speech signal has a level difference of −3 dB with respect to the calibration signal and the speech signal is to be presented at a level of 60 dB, the equipment is adjusted to yield a calibration signal level of 63 dB while playing the speech signal.

NOTE 1 Speech signals might be adjusted in their level in order to adjust their intelligibility. Hence, the L_{eq} of a speech signal cannot be equal to the L_{eq} of the calibration signal even when their nominal level difference is 0 dB.

NOTE 2 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.

4.6.3 Announcing signals

Announcing signals are optionally employed to be played in advance of a particular speech signal, e.g. to increase the attention. Their level might differ from the speech level, e.g., might be higher in order to be better understandable.

4.6.4 Competing sound signals

Competing sound signals are employed to determine the score (e.g., intelligibility) in a speech test at a particular speech-to-noise ratio (SNR), also referred to as speech test in noise. The relation (amplitude ratio or level difference in dB, respectively) of all competing sound signals to the signal used for the calibration shall be documented within the speech test. The nominal competing sound level is given by the level of the calibration signal when played at a nominal level difference of 0 dB.

NOTE Due to historical or other reasons, different competing sound signals may yield different L_{eq} values even when their nominal level difference is 0 dB.

4.7 Phonemic equivalence across test lists

The phoneme distribution shall be equal across test lists.

NOTE This does not require all phonemes be equally distributed within the test lists.

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

4.8 Perceptual equivalence across test lists

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

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. All speech recognition curves measured as given in 5.3 shall be within a range to be documented (see 5.5 k).

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

5 Validation of speech material recordings

5.1 General

For the validation, equipment fulfilling the requirements of IEC 60645-1 shall be used. Testing shall be performed monaurally for earphone or bone vibrator presentation. For sound field audiometry, the loudspeaker arrangement specified for the intended use of the speech material shall be applied. Only complete test lists shall be used.

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 of otologically normal persons of both sexes, between 18 years and 25 years inclusive, whose first language is the language of the speech material. 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 at or below 15 dB. If this criterion is not met, the minimum requirements for the pure-tone hearing threshold levels of test subjects shall be specified by the laboratory carrying out the validation.

NOTE The necessary minimum number of subjects can be determined according to Annex F.

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 of otologically normal children of the same age range as the speech material is intended to be used for.

To achieve the reference speech recognition curve, speech recognition measurements shall be performed at different presentation levels (or SNRs in case of a speech recognition test in competing sound). 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).

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

The reference speech recognition curve shall be given as a table which describes the relation between speech level (or SNR) and speech recognition score. The speech levels (or SNRs) which are required to achieve the speech recognition scores 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, and 90 % shall be specified. Usually, the measurements are carried out at predefined speech levels (or SNRs) so that the values for the required recognition scores are not directly available. Therefore, it shall be specified how the speech levels (or SNRs) for the recognition scores 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, and 90 % are calculated form the measurement data. One possibility is computing the median for each speech level (or SNR), and fit a discrimination function to these median values. In this case, the objective function and the parameters fitted shall be provided.

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.

NOTE Typical results are shown in Annex C.

5.3 Determination of perceptual equivalence of test lists

The perceptual equivalence of test lists of speech materials fulfilling the requirements specified in 4.8 shall be determined by performing speech recognition tests using a sufficiently large group 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 sound, the perceptual equivalence of test lists shall be determined using this competing sound. If the test is intended to be performed with different competing sounds, the perceptual equivalence of test lists shall be determined using each particular competing sound.

If the speech recognition test is intended for the determination of speech recognition threshold levels, the equality of test lists shall be specified for 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 for the speech recognition scores 50 %, 60 %, 70 %, 80 %, and 90 % when using different test lists with the same group of test subjects.

5.4 Determination of test-retest reliability

The test-retest reliability of a speech recognition test result is quantified by the 95 % confidence interval for the retest score for a given test score 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.

Since the 95 % confidence interval for the retest score depends on the speech recognition score for the test, test and retest scores shall be determined by performing speech recognition tests for different test scores using a sufficiently large group of otologically normal persons as described in 5.2.

NOTE The necessary minimum number of subjects can be determined according to Annex F.

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

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

If the test lists are equivalent and if each test item within a test lists has the same speech recognition score, the 95 % confidence interval for the retest can be estimated using theoretical assumptions. Sample data can be found in Annex G.

5.5 Documentation

Each recording shall include or be supplied with additional documentation containing the following information:

a) the content of the recorded speech material, e.g., vocabulary and test structure;

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

c) the reference speech recognition curves as determined 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 sound, if applicable);

d) the test-retest reliability of results as determined in 5.4;

e) the average frequency spectrum of the speech material;

f) the level differences of all signal types or signals, respectively, in accordance with 4.2 and 4.6;

g) the intervals between successive test items, if applicable;

h) the time and frequency characteristics and the level of any accompanying competing sound;

i) if the competing sound is presented in a gated condition (see B.4), the timing of the competing sound presentation;

j) the phoneme distribution of the test;

k) equivalence of test lists (see 4.8).

6 Requirements for speech audiometry

6.1 Audiometric equipment

Speech audiometry shall be performed using equipment fulfilling the requirements of IEC 60645-1. Testing may be performed either by earphones monaurally or binaurally, by bone vibrator, or in a sound field.

6.2 Ambient sound pressure levels in test room for speech audiometry

The ambient sound pressure levels in the test room shall not unduly affect the perception of 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.

6.3 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 loudspeaker that reproduces the speech signal 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 both the loudspeaker and the competing sound source(s) shall be identical to those used for the determination of the reference values.

The type of sound field (free, quasi-free or diffuse), as specified in ISO 8253-2, at least in the frequency range from 500 Hz to 4 kHz shall be stated.

7 Preparation and instruction of test subject

7.1 General

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

7.2 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.

7.3 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.

8 Subject's response mode

The speech material can be presented as an “open-set” test, in which the subject is not informed about the response alternatives, or as “closed-set” test in which the subject is informed about the response alternatives. The test subject's response is normally spoken but may be written or indicated by other means, e.g., pictures, human interface devices (HID). When the tester is not in the same room as the test subject or the tester has difficulties in understanding the subject’s response to the speech signal, which is spoken, a talk-back system, as described in IEC 60645-1, 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.

NOTE If understanding the response by the tester is not free of doubt, the following procedure is suggested prior to the speech test: The test subject is asked to read one or more sentences out loud in the one-to-one presence of the tester. The tester then satisfies himself or herself that they have heard and fully understood what the test person has read (see 15.h).

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 for non-adaptive speech tests.

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 equivalent to those employed during the determination of reference values.

9 Determination of speech detection threshold level

9.1 In this clause, an example for the determination of the speech detection threshold level is given. Alternative procedures shall be defined in the documentation of the speech test.

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

9.3 When determining a speech detection threshold level, 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.

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

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

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

9.7 Repeat step 9.6 until three responses have occurred at the same level out of a maximum of five ascents. This level is the speech detection threshold level.

9.8 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 lower than the level of the last response. Then repeat steps 9.5 to 9.7 inclusive.

9.9 Proceed with the other ear if required.

9.10 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 9.7 until at least two responses occur at the same level out of a maximum of three ascents.

10 Determination of speech recognition threshold level

10.1 General

In this clause, one example for the determination of speech recognition threshold levels is given. This example describes an adaptive procedure based on identifying the speech level where 50 % of the test items are correctly recognized.

NOTE The meaning of the term “correctly recognized”/”correctly responded” depends on the test procedure, see examples given in 10.6.

Alternative approaches shall be defined in the documentation of the speech test.

Reference values as well as the test-retest reliability in combination with the speech material shall be determined using a sufficiently large group of otologically normal persons as described in 5.2. For any reference values of the speech test, it shall be clearly specified which exact procedure and which settings (all relevant parameters) has been used to determine the respective values.

A test list shall contain enough test items to enable the procedure to adapt to the speech recognition threshold level.

NOTE A number of 20 test items are generally considered as sufficient.

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

10.3 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.

10.4 Reduce the speech level in steps of 5 dB, presenting one test item at each level, until the test subject no longer responds correctly to a test item.

10.5 Increase the level by 5 dB and present one test item.

10.6 Continue in accordance with 10.7.1 to 10.7.3 for single words or complete sentences as test items. For scoring based on keywords in sentences, follow steps 10.8.1 to 10.8.3.

10.7.1 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.

10.7.2 Repeat 10.7.1 for all remaining test items of the test list.

10.7.3 Discard the test item of 10.5 and the following test item, and average the presentation levels of the remaining items of the test list including the level which would be presented if the test list had one more item. This average is the speech recognition threshold level.

10.8.1 For the next 5 sentences: If the listener is able to repeat correctly more than half of the keywords, decrease the level by 2 dB. If the number of correctly repeated keywords is less than half, increase the level by 2 dB. If the number of correctly repeated keywords is exactly half, leave the speech level unchanged and present another test sentence.

10.8.2 For the remaining sentences: If the listener is able to repeat correctly more than half of the keywords, decrease the level by 1 dB. If the number of correctly repeated keywords is less than half, increase the level by 1 dB. If the number of correctly repeated keywords is exactly half, leave the speech level unchanged and present another test sentence.

10.8.3 Discard the measured speech levels for the sentence of 10.5 and the following sentence and average the presentation levels of the remaining sentences of the test list including the level which would be presented if the test list had one more item. This average is the speech recognition threshold level.

11 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 20 dB to 30 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 speech level(s) is governed by the particular purpose of the test. In the following, three examples are given for level choices. Alternative selections of speech levels shall be defined in the documentation of the speech test. The examples are:

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 levels lower than the level leading to this "roll-over effect".

b) To determine a speech recognition score at the most comfortable level for speech, the test speech 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 for speech.

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.

12 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 60645-1 specifies the requirements 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 as well as on the transducer type. 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:

$$L_M = L_t - 40\,\text{dB} + (L_{A\,min} - L_{B\,min}) \quad (1)$$

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.

13 Speech audiometry with competing sound

13.1 Type of competing sound

A frequency-weighted unmodulated random noise in accordance with IEC 60645-1 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.

NOTE 1 Examples of competing sounds are given in Annex B.

NOTE 2 It is recognized that competing sounds using recordings of human speech can also be used, such as multi-talker 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.

13.2 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.

13.3 Speech and competing sound levels

The speech level and competing sound level shall be measured as specified in the documentation of the speech test. If only one measurement can be made, 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.

NOTE If the competing sound is not clearly audible to the subject, the result is not a valid SRT in noise and thus can be misleading.

13.4 Test procedure

13.4.1 General

The test may be performed either to determine a speech recognition score at one or more fixed SNRs or to determine the SNR at the speech recognition threshold level.

13.4.2 Speech recognition score at fixed SNR

13.4.2.1 Set the audiometric equipment to the required speech level.

13.4.2.2 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.

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

13.4.3 Speech recognition threshold level with a competing sound

The example procedure for determining the speech recognition threshold level with a competing sound is the same as the example procedure described for determination of speech recognition threshold level for speech in quiet in Clause 10 with the following exceptions:

— As to 10.3: 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. 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.

— In the measurement phase the competing sound shall be presented at an audible level. The recommended level for the competing sound is 65 dB. The recommended SNR for the first presentation is 0 dB.

— As to 10.7.3 and 10.8.3: Average the SNRs instead of the presentation levels.

Alternative approaches shall be defined in the documentation of the speech test.

Reference values as well as the test-retest reliability in combination with the speech material shall be determined using a sufficiently large group of otologically normal persons as described in 5.2. For any reference values of the speech test, it shall be clearly specified which exact procedure and which settings (all relevant parameters) has been used to determine the respective values.

In the procedure described, decreasing the level of the speech signal may be replaced by increasing the level of the competing sound. The test result may be influenced by the alternative chosen.

14 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 SNR, in decibels, along the abscissa. The scale ratio should be 20 % corresponding to 10 dB. The type of speech material, any competing sound, and the spatial positions of speech and competing sound sources used shall be reported.

A speech audiogram should contain the corresponding reference speech recognition curve as well as its test-retest reliability (expressed as ±2 standard deviations) for the actual speech material used.

Symbols for a graphical representation of the results are given in Annex C.

15 Measurement uncertainty

The uncertainty of speech recognition scores determined in accordance with any of the procedures specified in this document 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, experience, dialect, hearing ability and lexical skills of the tester;

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

h) the reliability of response detection and scoring process (see Clause 8)

i) the use of non-optimized contralateral masking noise (see Clause 12)

j) the use of non-optimized competing sound.

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 document shall be evaluated in accordance 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.

16 Maintenance and calibration of equipment

16.1 General

Correct calibration of audiometric equipment is essential 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 (see 16.3);

— stage B: periodic electroacoustic tests;

— stage C: basic calibration tests.

16.2 Stage A: Routine checking and listening tests

16.2.1 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 16.2.2 to 16.2.9.

16.2.2 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.

16.2.3 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 to use the transducers calibrated in conjunction with the audiometer.

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

16.2.5 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.

16.2.6 Listen to the speech test material at a higher level (e.g., at a speech level 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.

16.2.7 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.

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

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

16.3 Stage B: Periodic electroacoustic tests

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

a) the frequency responses for the equipment including all appropriate output transducers as specified in IEC 60645-1;

b) the output levels of the transducers as specified in IEC 60645-1;

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 60645-1;

e) the masking noise levels in accordance with IEC 60645-1;

f) the headband force of the transducers.

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

16.4 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 60645-1 and this document.

16.5 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 months 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.

Annex A
(informative)

Example of speech materials

A.1 General

Speech materials can be divided into different classes. This annex provides references that describe 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 [13] to [16] and [31].

A.3 Examples of closed-set materials

Examples of closed-set materials are given in References [17] to [19] and [30].

A.4 Examples of open-set sentence materials

Examples of open-set sentence materials are given in References [20] to [25] and [28].

A.5 Examples of closed-set sentence materials

Examples of closed-set sentence materials are given in References [26] to [28].

Annex B
(informative)

Examples of competing sound

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 competing sound shows no amplitude modulation or only a very small amount of amplitude modulation. 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 speech-simulating noise, e.g., the CCITT noise (ITU-T Recommendation G.227[9]; telephone frequency characteristic) or the ICRA1 noise (see Reference [32]) 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 (see 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 (see 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 [34]. In this noise the amplitude modulations are the same for all frequencies.

B.4 Babble noises

This kind of competing 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 14).

Annex C
(informative)

Typical results

C.1 Typical reference speech recognition curves

As an example, data for the Göttingen sentences (see Reference [21]) 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

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

where

L is the speech level (see 3.16) in quiet or SNR in competing noise;

L_{50} is the speech recognition threshold level (see 3.21) in dB (50 % correct);

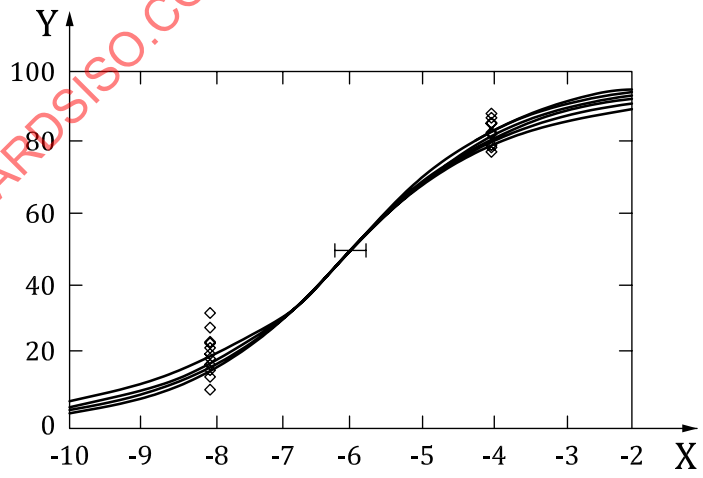
NOTE For tests without competing noise: L_{50} = 20 dB. For tests with speech simulating unmodulated noise, the SNR is L_{50} = −6 dB.

S is the slope of the recognition curve in percentage points/dB at L_{50};

S_S 11 percentage points/dB in silence;

S_N 19 percentage points/dB in noise.

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



Key

X SNR, expressed in dB

Y proportion of correct words, expressed in %

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

NOTE The symbols at −4 dB 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 level. The test lists were optimized for the same speech recognition threshold level.

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	[21]	Words	20 dB
Open-set sentences in noise	[21]	Words	SNR: −6 dB
Open-set sentences in noise	[20]	Words	SNR: −5 dB
Open-set sentences in noise	[20]	Whole sentences	SNR: −3 dB
Open-set sentences in noise	[24]	Whole sentences	SNR: −6 dB
Open-set sentences in noise	[23]	Whole sentences	SNR: −3 dB
Open-set words in silence	[14]	Words	28 dB
Open-set words in silence	[31]	Words	23 dB
Open-set words in silence	[14]	Phonemes	24 dB
Closed-set sentences in noise	[28]	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 (see Annex F) 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	[21]	Words	2 dB
Open-set sentences in noise, 10 sentences	[20]	Whole sentences	0,9 dB
Open-set sentences in noise, 10 sentences	[22]	Whole sentences	1,2 dB
Open-set words in silence, 50 words	[14]	Words	10 %
Open-set words in silence, 50 words	[14]	Phonemes	8 %
Closed-set sentences in noise, 20 sentences	[28]	Words	1 dB

C.4 Speech audiograms

Speech reception scores may be presented in a speech audiogram (see Clause 14). The symbols given in Table C.3 are suggested.

Table C.3 — Symbols for the graphical representation of speech audiometry results

	Right	Binaural	Left
Speech test in quiet			
Speech test with contralateral masking			
Speech test with competing sound			
Speech test with competing sound with contralateral masking			
Speech test with competing sound – sound field, unaided			
Aided speech test with competing sound – sound field, device 1			
Aided speech test with competing sound – sound field, device 2			
Aided speech test with competing sound – sound field, device 2 and indication “a”, “b”, “c” for test materials			

The basic symbols for right ear, left ear or binaural results are based on the symbols in ISO 8253-2, which are used for speech scores. Any kind of binaural earphone presentation (diotic or dichotic) is indicated as binaural. Monaural sound field presentation is achieved by preventing usage of the other ear by suitable means.

Modifiers (alone or in combination) are used to indicate other details of particular measurements.

— A filled triangle on top of the symbol indicates that the measurement has been performed with contralateral masking. It is assumed that this does not apply for binaural measurements.

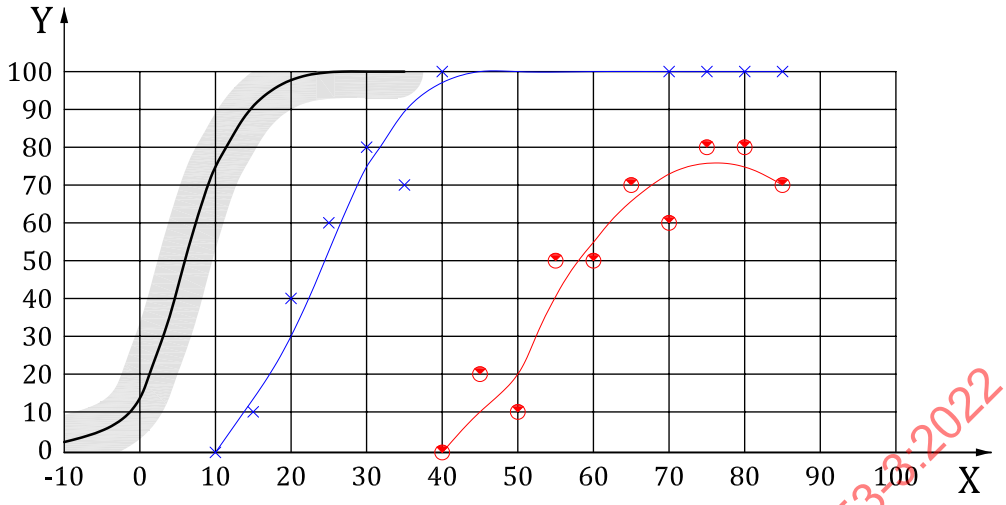
— A horizontal line below the symbol indicates a measurement with a competing sound.

— A horizontal line on top of the symbol indicates measurements performed in a sound field.

— A diamond around the symbol indicates measurements using a hearing system (aided measurements). If tests are being conducted with more than one device, a number on the right side of the symbol indicates the device used from the second device onwards.

— A letter on the left side of the symbol indicates more details if applying, for instance the specific speech material used in the test. A legend must be given in this case.

Figure C.2 presents an example of a speech audiogram for a typical monosyllabic-word test material. The grey line and shaded area show the reference curve and the standard deviation. The measurement for the left side (blue symbols and blue curve) shows typical results for a mild-to-moderate hearing loss. The measurement for the right side (red symbols and red curve) shows typical results for a more severe hearing loss. The difference between the highest score yielded and 100 % is called discrimination loss.



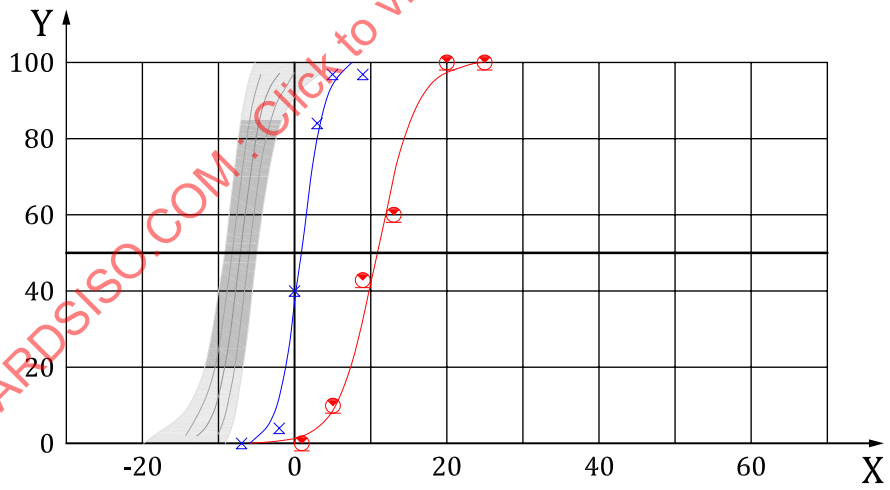
Key

X speech level, expressed in dB

Y speech recognition score, expressed in %

Figure C.2 — Example of a speech audiogram for a typical monosyllabic-word test material

Figure C.3 presents an example of a speech audiogram for a typical sentence test material. The grey line and shaded area show the reference curve and the standard deviation. The measurement for the left side (blue symbols and blue curve) shows typical results for a mild-to-moderate hearing loss. The measurement for the right side (red symbols and red curve) shows typical results for a more severe hearing loss.



Key

X SNR, expressed in dB

Y speech recognition score, expressed in %

Figure C.3 — Example of a speech audiogram for a typical sentence test material

Annex D
(informative)

Optimization of perceptual equivalence of test lists

The perceptual equivalence of test lists may be optimized in the following way[21]:

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 should not yield perceivable loudness differences between the particular test items within the test material (e.g. obvious loudness differences between the particular words of a test sentence).

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 document 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 document 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 Formula (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 document (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 the lack of qualification and experience of the tester (E.3.6);

δ_{su} is an input quantity to allow for any uncertainties due to the lack of co-operation and unreliable 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.

It is possible that, for other measurements, Formula (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 speech recognition 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 60645-1 for a Class 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 60645-1 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 $\sqrt{3}$.

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 $2{,}5/\sqrt{3}$ dB.

The two contributions result in an approximate overall standard uncertainty, e.g., for air conduction, given by Formula (E.2):

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

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.

E.3.5 Masking and competing sound, 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 (see E.3.2).

E.3.6 Impact of the tester, δ_{te}

For a qualified tester with appropriate experience and skills, the contribution to uncertainty due to personal judgements may be considered to be included in the standard uncertainty for repeated measurements (see 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 (see 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 recordings 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 6.2) 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 level, 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 10.2 to 10.6. 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 speech recognition 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, see Table E.1.

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

Quantity	Estimate of quantity value dB	Standard uncertainty u_i dB	Probability distribution	Sensitivity coefficient c_i	Uncertainty contribution $c_i u_i$ dB
L'_{HT}	$L'_{HT,est}$	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 Formula (E.3):

$$u = \sqrt{\sum_{i=1}^{10} u_i^2} \tag{E.3}$$

ISO/IEC Guide 98-3 requires an expanded uncertainty, U, to be specified, such that the interval $[L_{HT} - U, L_{HT} + U]$ 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 contributions 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 of quantity value 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	2,5	normal	1	2,5
δ_{mth}	0	2	normal	1	2

Combined standard uncertainty: u = 4,7 dB.

Expanded measurement uncertainty for 95 % coverage probability, rounded to the nearest decibel: U = 9 dB.

Annex F
(informative)

Determination of minimum number of subjects

A speech test can be modelled as a Bernoulli process (Hagerman, 1976[35]). The recognition score obtained from one test list is considered as a random variable. If each test item within a test list of n items has the same probability p (in percent) for recognition, the expectation value for the recognition score is p, and the standard deviation for the recognition score of this test list is:

$$\sigma = \sqrt{\frac{p(100-p)}{n}} \tag{F.1}$$

The standard deviation has its maximum at p = 50 %. Let the recognition score of the test list be obtained by m subjects. The expectation value of the mean score is again p, the standard deviation of the mean score is $\frac{\sigma}{\sqrt{m}}$. The 95 % confidence interval for the mean score can be approximated by $[p-c_{95}, p+c_{95}]$ with

$$c_{95} = 1{,}96 \cdot \sqrt{\frac{p(100-p)}{n \cdot m}} \tag{F.2}$$

Formula (F.2) can be solved for m to estimate the number of subjects required to determine the speech recognition score for a specific test list with a given uncertainty c_{95}:

$$m = \frac{p(100-p) \cdot 1{,}96^2}{n \cdot c_{95}^2} \tag{F.3}$$

Example 1 respective 5.2 One test list has $n = 20$ items. The target speech recognition score shall be $p = 50\,\%$. The uncertainty shall be $c_{95} = 4{,}5\,\%$, i.e., it is assumed that the mean score of m subjects lies with probability 95 % in the interval $[45{,}5\,\%, 54{,}5\,\%]$. Formula (F.3) then gives:

$$m = \frac{50\,\%(100\,\%-50\,\%) \cdot 1{,}96^2}{20 \cdot (4{,}5\,\%)^2} = 23{,}7$$

As a result, the number of subjects required is 24.

Example 2 respective 5.3 Each test list has $n = 20$ items. Two test lists with a difference in the speech recognition scores of more than 4,5 % at a speech recognition score of 50 % are considered to be not equivalent. It is assumed that the two test lists have the same score variance σ^2 / m and that the variance σ_d^2 of the difference between the scores of the two test lists is the sum of the variances of each test list[37]: $\sigma_d^2 = 2\sigma^2 / m$. If the two lists had the same recognition probability p, an observed difference between the two averaged scores over m subjects would be with probability 95 % in the interval $[-1{,}96\,\sigma_d, 1{,}96\,\sigma_d]$. This can be regarded as the range of uncertainty of the measured score difference. If it is as an example required that this uncertainty be equal or less than 4,5 %, it follows that at most $1{,}96 * \sigma_d = 4{,}5\,\%$. Consequently, the uncertainty of each of the two test list scores shall be $c_{95} = \frac{1{,}96\sigma}{\sqrt{m}} = \frac{1{,}96\sigma_d}{\sqrt{2}} = \frac{4{,}5\,\%}{\sqrt{2}} = 3{,}2\,\%$, i.e., it is assumed that the mean score of m subjects lies with probability 95 % in the interval $[46{,}8\,\%, 53{,}2\,\%]$. Formula (F.3) then gives: