FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/FDIS
8253-3**

ISO/TC **43**

Secretariat: **DIN**

Voting begins on:
2021-10-22

Voting terminates on:
2021-12-17

Acoustics — Audiometric test methods —

Part 3: Speech audiometry

Acoustique — Méthodes d'essais audiométriques —

Partie 3: Audiométrie vocale

ISO/CEN PARALLEL PROCESSING

Reference number
ISO/FDIS 8253-3:2021(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

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).