TECHNICAL SPECIFICATION

# ISO/TS 9241-126

First edition
2019-12

# Ergonomics of human-system interaction —

## Part 126: Guidance on the presentation of auditory information



Reference number
ISO/TS 9241-126:2019(E)



ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
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 of 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 www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 159, *Ergonomics*, Subcommittee SC 4, *Ergonomics of human-system interaction*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

# Introduction

Interactive systems have used auditory information as primary sources and, more commonly, to enhance interactions in primarily visual systems since about the time the digital computer was invented. Even before that, one can consider telephone conversations to be a type of interactive system where information is presented aurally. The way participants in a telephone conversation present information to one another is largely based on learned and subconscious conventions. In the digital realm, sound has been used in increasingly symbolic ways, straying out of necessity from the realistic representations of objects and actions in the physical world.

Unlike visual displays, the information conveyed through sound is typically sparse, and it seems more difficult to intuitively understand what makes a sound scheme usable, whether in its design or its presentation. Since the main communication medium of modern interactive software is the visual display, it is not surprising that a need was first identified for guidelines relating to presentation of visual information. However, designers developing for both visual and auditory information would benefit from guidelines rooted in scientific discovery, that lead to clear and usable communication of information, while not inhibiting the creative process.

This document focuses on guidelines for presentation of information in the auditory modality.

The success of this document will depend on its positive impact on usability of future systems presenting auditory information using these guidelines. However, this document is not necessarily designed to be used by the direct beneficiaries of this desired usability increase. Instead, it serves the following types of users:

a) the user interface designer, who will apply these guidelines during the development process;

b) the buyer, who will reference this document during the product procurement process, and whose end users will gain from the potential benefits provided by the guidelines;

c) those responsible for ensuring products meet the recommendations in this document;

d) designers of auditory interface development tools to be used by auditory interface designers;

e) writers of software industry standard guides to be used by auditory interface designers.

# Ergonomics of human-system interaction —

# Part 126:
# Guidance on the presentation of auditory information

## 1 Scope

This document provides guidance for the auditory presentation of information controlled by software, irrespective of the device. It includes specific properties such as the syntactic or semantic aspects of information, e.g. coding techniques, and gives provisions for the organization of information taking account of human perception and memory capabilities.

This document does not address the hardware issues of the transmission and the production of auditory information.

NOTE 1 Volume is dependent on hardware and thus cannot always be absolutely controlled by software. Environmental conditions can also affect the ability for sounds to be perceived, which can be beyond the ability of the software to take into account.

This document does not apply to auditory alarms, warnings or other safety-related uses of auditory information.

NOTE 2 Safety-related uses of auditory presentation of information are covered in various domain specific standards, such as ISO 7731:2003 which deals with auditory danger signals for public and work areas, and IEC 60601-1-8:2006 which provides very specific requirements for auditory alarms for medical devices.

While this document applies to the presentation of all non-safety-related information, it does not include application domain specific guidance (e.g., audio instructions for consumer products).

This document can be utilized throughout the design process (e.g. as specification and guidance for designers during design or as a basis for heuristic evaluation). Its provisions for the presentation of information depend on the auditory design approach, the task, the user, the environment and the single or multiple technologies that can be used for presenting the information. Consequently, this document cannot be applied without knowledge of the context of use. It is not intended to be used as a prescriptive set of rules to be applied in its entirety but rather assumes that the designer has proper information available concerning task and user requirements and understands the use of available technology.

This document does not address visual or tactile/haptic presentation of information or modality shifting for the presentation of auditory information in other modalities.

NOTE 3 ISO 9241-112 provides high-level ergonomic guidance that applies to all modalities.



## 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 9241-171, *Ergonomics of human-system interaction — Part 171: Guidance on software accessibility*

## 3 Terms and definitions

For the purposes of this document, the following terms and definitions 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 http://www.electropedia.org/

**3.1**
**pitch**
perception of the highness or lowness of a sound's *frequency* (3.2)

**3.2**
**frequency**
number of complete oscillations per unit time of a periodic waveform

**3.3**
**fundamental frequency**
lowest *frequency* (3.2) component of a periodic waveform

**3.4**
**overtone**
higher *tone* (3.18) produced simultaneously with a *fundamental frequency* (3.3) and that, with the fundamental frequency, comprise a complex tone

**3.5**
**harmonic**
*overtone* (3.4) that is an integer multiple of the *fundamental frequency* (3.3)

**3.6**
**harmonicity**
degree to which a sound's *overtones* (3.4) are composed of integer multiples of its *fundamental frequency* (3.3)

Note 1 to entry: A sound with high harmonicity is called "harmonic" while a sound with low harmonicity is called "inharmonic".

Note 2 to entry: The degree to which a sound's overtones depart from being integer multiples of its fundamental frequency is called "inharmonicity".

**3.7**
**pure tone**
*tone* (3.18) composed of a single *frequency* (3.2) with no *harmonics* (3.5)

**3.8**
**intensity**
power per unit area of a sound wave in a direction perpendicular to that area

**3.9**
**loudness**
perception of a sound's *intensity* (3.8)

**3.10**
**masking threshold**
sound pressure level below which a signal is not audible when in the presence of another specified sound (i.e. a masking signal or masker)

Note 1 to entry: The masking threshold depends on the sound's *frequency* (3.2), the type of masker, and the kind of sound being masked.

**3.11**
**timbre**
quality given to a sound by its *overtones* (3.4)

**3.12**
**reverberation**
**resonance**
**echo**
prolongation of a sound through reflection off objects or structures in the sound environment

**3.13**
**interaural**
relating to sound reception and perception by each ear considered separately

**3.14**
**monaural**
relating to sound emanating from a single sound source or presented equally from all sound sources

**3.15**
**rhythm**
strong, regular, repeated pattern of sound

**3.16**
**accent**
increased stress or emphasis on a particular sound within a sequence, usually by increasing its *intensity* (3.8) or duration relative to other sounds in the sequence

**3.17**
**earcon**
**auditory icon**
brief, distinctive sound used to represent a specific object or event

**3.18**
**tone**
discrete musical sound

**3.19**
**prosody**
rhythmic and intonational aspect of language

**3.20**
**lateral sound localization**
ability of a user to approximate the location where a sound originated

**3.21**
**accessibility**
extent to which products, systems, services, environments and facilities can be used by people from a population with the widest range of user needs, characteristics and capabilities to achieve identified goals in identified contexts of use

Note 1 to entry: Context of use includes direct use or use supported by assistive technologies.

[SOURCE: ISO 9241-112:2017, 3.15]



# 4 Application

## 4.1 Accessibility

Auditory presentation of information shall be in accordance with ISO 9241-171, which gives specific requirements and recommendations for the presentation of information in the auditory modality and the presentation of the same information in different modalities.

NOTE ISO 9241-171:2008, 10.6, contains significant guidance on the controllability and accessibility of audio output.

### 4.2 Applying the recommendations in this document

It is recognized that different clauses in this document apply in different situations:

— Clause 5 contains recommendations on appropriate usage of auditory presentation of information;

— Clause 6 contains general recommendations on the auditory presentation of information;

— Clause 7 contains recommendations on auditory dimensions;

— Clause 8 contains recommendations on speech;

— Clause 9 contains recommendations on earcons;

— Clause 10 contains recommendations on auditory coding.

Each individual recommendation should be evaluated for its applicability and, if judged to be applicable, should be implemented, unless there is evidence that doing so would cause deviation from the design objectives, or would result in an overall degradation in usability. In judging whether applicable recommendations have been met, evaluators should evaluate the product or observe representative users of the product in the context of accomplishing the user's task.

## 5 Appropriate usage

### 5.1 Visual system unavailability

If the user's visual system is either unavailable or overburdened, then consider presenting information aurally unless there is reason to prefer tactile communication.

EXAMPLE 1 A blind user's visual system is unavailable, so the user opts for an auditory presentation.

EXAMPLE 2 In an application with a dense visual display of information, additional information is provided by the auditory channel to possibly increase the user's task performance efficiency.

### 5.2 Inadequate or excessive light

If the ability to perceive visually is limited by excessive light or darkness, including the need to maintain dark adaptation, consider presenting information aurally.

### 5.3 Inability to maintain visual attention

If the task prevents the user from maintaining focus on a visual display, then consider presenting information aurally.

EXAMPLE An auditory presentation used for a shop floor inventory application because workers using the application are moving around the shop floor and also using both hands much of the time.

### 5.4 Message brevity and simplicity

Auditory messages should be considered when:

1) messages can be short and simple;

2) other means of communication are not practical; and/or

3) aural presentation is more natural for the information to be presented than other modalities of communication.



 

### 5.5 Event-based messages

If a message relates to or signifies an event occurring at the time of presentation, then consider presenting the message aurally.

### 5.6 Messages requiring immediate action

If a message requires immediate action, consider presenting the information aurally.

### 5.7 Messages requiring verbal response

If a message requires a verbal response, consider presenting the information aurally.

### 5.8 Permanence of message relevance

If a message needs to be referred to later, aural messages should not be used alone (i.e. a redundant presentation of the message in a more permanent manner).

### 5.9 Attention capture

When the attention is required of a user who can be focusing on other tasks, consider using sound to capture the user's attention.

EXAMPLE A user looking away from a visual display is alerted by an audible tone to look at new information available on the display.

### 5.10 Muting sound

When a sound is used to capture the user's attention, the user should have the option of silencing it.

## 6 Auditory presentation

### 6.1 General

#### 6.1.1 Auditory channel overload

Auditory presentation of information should be used sparingly so as not to overwhelm the user's auditory channel.

NOTE The auditory channel does not have as much bandwidth as the visual channel, so care needs to be taken not to present too much information aurally.

#### 6.1.2 Information conciseness

Auditory presentation should be limited to only necessary information unless redundant or superfluous information improves detectability, discriminability, or learnability.

#### 6.1.3 Avoid extreme dimensions

Extreme values in any auditory dimension should be avoided whenever possible.

NOTE 1 Extremely loud sounds can be very unpleasant. Quiet sounds and extremely high and low frequencies are difficult to detect.

NOTE 2 Frequent exposure to sounds of greater than 85 dB in intensity can damage hearing.

NOTE 3 People with normal hearing are maximally sensitive to sounds in the frequency range of speech (i.e. 2 000 Hz–5 000 Hz).

### 6.1.4 Monaural option

If a stereo or multi-channel sound field is used, users should have the option whenever possible to have the sound presented monaurally.

NOTE Users with reduced hearing in one ear can miss information if it is presented in stereo.

### 6.1.5 Ease of learning

Whenever possible, sounds should be used that are commonly associated with what is being represented or whose properties are analogous to some property being represented.

EXAMPLE 1 A sound with rising frequency is used to represent filling up a container.

EXAMPLE 2 In a home healthcare application, the sounds of pills rattling in a pill bottle is used to signal the user that it is time to take medication.

### 6.1.6 Signal invariance

The same signal should be used to present the same information at all times.

### 6.1.7 Signal similarity

Sounds that represent reasonably different information should subjectively sound reasonably dissimilar.

NOTE This is derived from the Gestalt law of similarity.

### 6.1.8 Two-stage signal usage

a) If the information being presented is complex, an initial signal to gain the user's attention should be presented before presenting the information.

b) If the initial signal is only used to alert users, then the same signal should be used as the alerting signal throughout the application.

EXAMPLE If the user's attention is not focused on listening, the user can miss the first part of the signal or message. Presenting an alerting initial signal prepares the user to listen.

### 6.1.9 Logical sequencing

Pieces of information should be presented in logical order, particularly if a sequence is required for data entry relating to this information.

EXAMPLE If a sequence of audio signals represents a set of actions that needs to be taken in a particular order, present them in that order.

### 6.1.10 Supporting short term memory

The number of items of information presented at one time should be limited in order to support memorability.

EXAMPLE A banking interactive voice response system limits the number of choices at one time to between 3 and 5 items.

NOTE Items presented earlier in the sequence are generally more memorable due to the primacy effect.



 

#### 6.1.11 Cross-cultural sounds

If sounds need to be used in different countries or cultures, or will be presented to individuals speaking different languages, then the sounds should be culturally appropriate.

NOTE Some sounds can have specific cultural meanings associated with them, and using them for a different meaning can be confusing or offensive.

#### 6.1.12 Representative sounds

If the information is auditory in nature, or if a natural auditory analogue for the information exists, then that representative sound should be used to present that information.

EXAMPLE A heart monitor has an auditory display of heart rate that sounds like a beating heart.

#### 6.1.13 Representative sound usage

If the task involves distinguishing between very similar sounds, then additional information should be presented to support the distinction.

EXAMPLE If users need to distinguish between the calls of three birds, all of which sound very similar, then providing additional information such as the names of the birds would be appropriate for the task.

NOTE This can be addressed with redundant presentation of information visually.

### 6.2 Detectability in noisy environments

#### 6.2.1 Signal distinctiveness

Auditory signals should be clearly different from any other simultaneously presented audio output.

#### 6.2.2 Change signal frequency

The signal frequency, when possible, should lie within a spectral range in which the noise intensity is low.

#### 6.2.3 Signal duration in noisy environments

Signals that need to be detected in noisy environments should be presented for at least 500 ms.

## 7 Auditory dimensions

### 7.1 Frequency

#### 7.1.1 Absolute pitch identification

Frequency should not be used on its own to convey information when an absolute judgment is needed.

#### 7.1.2 Temporal proximity for pitch comparison

When two frequencies are presented for direct comparison, the two tones should be presented in quick succession.

#### 7.1.3 Similarity for pitch comparison

When frequencies are presented for direct comparison, the timbres and intensity levels of the tones should be as similar as possible.

NOTE Perceived pitch tends to drop with increasing intensity below 1 kHz and tends to rise when above 1 kHz.

#### 7.1.4 Minimum difference for pitch comparison

When frequencies are presented for direct comparison, there should be at least a 3 Hz difference between frequencies below 500 Hz and at least a 0,6 % difference between frequencies above 1 kHz.

#### 7.1.5 Frequency range for pitch identification

Frequencies below 80 Hz should not be used for any task requiring pitch identification.

#### 7.1.6 Frequency range for most accurate pitch discrimination

When accurate pitch discrimination is required, frequencies of between 1 kHz and 2 kHz should be used.

#### 7.1.7 Harmonicity for pitch identification

Inharmonic tones should not be used for any task requiring pitch identification.

EXAMPLE Users are unlikely to be able to identify the pitch of a click or a percussive drum sample. Truly inharmonic sounds do not have a discernible pitch.

#### 7.1.8 Harmonic tones for pitch discrimination tasks

a) If possible, harmonic tones, rather than pure tones, should be used for pitch discrimination tasks.

b) Inharmonic tones should not be used for pitch discrimination tasks.

NOTE Most people have great difficulty identifying the pitch of inharmonic sounds, and there is some evidence that pitch discrimination is somewhat better for harmonic tones than for pure tones.

#### 7.1.9 Long-distance sound

If the signal must travel over a large distance, then frequencies less than 1 kHz should be used.

#### 7.1.10 Obstructed sound

If the signal must travel past major obstacles, then frequencies less than 500 Hz should be used.

EXAMPLE If the sound needs to be heard on the other side of a closed door, use frequencies less than 500 Hz.

### 7.2 Intensity

#### 7.2.1 Avoid distortion

Increasing audio volume should not distort sounds.

#### 7.2.2 Separate signal control

Users should be able to control levels of particular types of sound separately, especially in the case that audio is used only to supplement visual information.

EXAMPLE A listener mutes background music and increases the audio volume of messages from the computer application that the listener is using.



 

### 7.2.3 Loudness cues about the environment

Signal intensity may be used to convey characteristics and conditions of the environment where applicable.

NOTE Different relative levels of loudness can suggest the presence of obstacles or convey information about object composition, distance, and sound absorption.

EXAMPLE Sound intensity is employed as a proximity "code" for objects in an application. Objects closer to the user are represented by sounds of greater intensity than objects that are farther away from the user.

### 7.2.4 Minimal signal intensity

Audio signals should be at least 40 dB and at least 6 dB above the masking threshold.

NOTE A higher minimum signal intensity applies in some situations, such as: where the users are likely to be wearing hearing protection or when the target users of the software application are elderly persons (for example, an app for use by elderly persons living alone to call for help in medical or other emergencies).

### 7.2.5 Signal intensity for rapid response

Audio signals requiring rapid response should be at least 16 dB above the masking threshold.

### 7.2.6 Relative maximal signal intensity

Audio signals should be no more than 30 dB above the masking threshold.

### 7.2.7 Absolute maximal signal loudness

Audio signals should not be more than 115 dB.

## 7.3 Timbre

### 7.3.1 Timbre discrimination

Different timbres used for different purposes should be subjectively easy to discriminate.

### 7.3.2 Musical instrument timbres

Wherever possible, musical instrument timbres should be used instead of pure tones.

## 7.4 Localization

### 7.4.1 Guidance related to lateral sound localization

**7.4.1.1** Use interaural time and intensity differences to convey information about lateral sound localization relative to the user.

**7.4.1.2** Time differences in the lower frequencies should be used to facilitate lateral sound localization.

**7.4.1.3** Echoes and reverberation effects should not be used to enhance lateral sound localization.

### 7.4.2 Minimum angle for position discrimination

If the task requires discrimination of two sound source positions, the angle between the two sources should be at least 2 degrees. If the sound sources are different, the minimum angle should be increased.

NOTE See also 6.1.4.

### 7.4.3 Frequency selection for sound localization

If users are required to localize sound, frequencies less than 1,5 kHz and/or greater than 5 kHz should be present in the sound source spectrum. Sounds with frequency components spread across the entire spectrum should be used to facilitate sound localization.

NOTE See also 6.1.4.

### 7.4.4 Minimum distance for front/back position identification

Identification of front/back sound position should not be required for distances of less than 50 cm.

### 7.4.5 Guidance related to distance judgment

**7.4.5.1** Relative loudness should be used to convey information about the distance of a particular sound source.

**7.4.5.2** Relative reverberation should be used to facilitate distance judgment.

NOTE Reverberation can inhibit localization and impede speech intelligibility.

**7.4.5.3** Sound sources located laterally relative to the user should be used for accurate distance judgment.

**7.4.5.4** If users are required to estimate distances of less than 1 m, interaural time and intensity differences and low-frequency content should be used to facilitate distance judgment.

### 7.4.6 Doppler effect

Rising and falling pitch should be used to convey motion of a sound source toward and away from the user. The amount of pitch change should be greater when the sound source is closer to the user.

### 7.4.7 Guidance related to elevation angle judgment

**7.4.7.1** If the user is required to localize the angle of elevation of a sound source, then sounds should be used that are directly in front of the user.

**7.4.7.2** If users are required to localize the angle of elevation of a low-frequency sound source that is not directly in front of them, then interaural time differences should be used to facilitate localization.

**7.4.7.3** If users are required to localize the angle of elevation of a high-frequency sound source that is not directly in front of them, then interaural intensity differences should be used to facilitate localization.



 

## 7.5 Duration

### 7.5.1 Minimum signal duration

For a signal to be accurately and reliably detected, it should have duration of at least 100 ms.

NOTE In high-noise environments, it is important to increase signal duration to at least 500 ms (see also 6.2.3).

### 7.5.2 Adjustable signal duration

The user should be enabled to set a minimum/maximum duration for signals, as long as that is greater than specified in 7.5.1.

## 7.6 Rhythm and accent usage

Rhythm and note accents should be used to ease and improve recognizability of sounds.

## 7.7 Timing

### 7.7.1 Event-based sounds

Event-based sounds should be presented along with some other event cue in another modality that has meaning so that the user can learn and remember the association.

EXAMPLE An audio sound (e.g. buzzer) indicates that clothes in a dryer are dry, and is accompanied by a visual display indicating that the dry cycle is completed.

### 7.7.2 Sounds to guide rhythmic movements

Sounds should be used to denote intervals for repeated rhythmic movements.

EXAMPLE A repetitive sound can be presented at regular intervals to guide finger tapping.

### 7.7.3 Concurrent auditory streams

a) If users are required to differentiate between two or more concurrent auditory streams, then differences in frequency, localization information, and onset and offset timings should be increased to facilitate differentiability.

b) Tones that share many of the same overtone frequencies should not be used in these situations.

NOTE If auditory patterns are being presented at the same time and the user needs to pay attention to both simultaneously, it is important for the user to perceive them as separate auditory streams rather than as one.

# 8 Speech

## 8.1 Appropriate usage

### 8.1.1 Complex messages

If an auditory message is complex and can be more easily communicated and understood as language, then speech should be used.

### 8.1.2 Source identification

If the user needs to be able to identify the person who is the source of the message, then speech should be used.

#### 8.1.3 Untrained users

If it is not possible or practical to supply users with special training in coded signals, then speech should be used.

#### 8.1.4 Stressful context of use

If a stressful environment or task can cause the user to forget the meaning of a coded message, then speech should be used.

#### 8.1.5 Rapid information exchange

If the task requires a rapid, two-way exchange of information, then speech should be considered.

### 8.2 General speech presentation

#### 8.2.1 Eliminate non-relevant speech

A speech message should contain as little inessential information as possible.

#### 8.2.2 Audio volume consistency

Prompts and announcements should be recorded and presented at a consistent audio volume throughout the application.

#### 8.2.3 Monophonic speech presentation

Speech messages should be presented monophonically.

#### 8.2.4 Present messages serially

a) No more than one speech message should be presented at the same time.

b) Relative importance should be used to prioritize the order of presenting messages.

NOTE Because of the complexity of speech messages and the density of information, it is especially difficult to pay attention to two such messages simultaneously.

#### 8.2.5 Information brevity

If the information needs to be recalled or remembered, no more than four pieces of information should be presented in a single message to avoid overloading short-term memory.

#### 8.2.6 Message replay

a) Users should be provided with a means of repeating messages.

b) If technically feasible, users should be provided with a means of repeating portions of long messages.

#### 8.2.7 Option phrase sequencing

Prompts that require user response from a list of options should be structured so that the option precedes the action.

NOTE This helps the user to pay closer attention when relevant information is about to be presented.

EXAMPLE Use "For y, press x" instead of "Press x for y."



 

#### 8.2.8 Grouping information

When multiple pieces of information are presented in one message, pauses should be inserted between phrases to group information.

NOTE Inserting pauses between pieces of distinct information to delimit options helps to to identify the individual pieces of information and to clarify their boundaries.

### 8.3 Language

#### 8.3.1 Language and dialect control

Where possible, the user should be able to select a preferred language or dialect.

#### 8.3.2 Avoid dialect-specific expressions

If selection of a preferred dialect is not possible or practical, then terms or expressions unique to a particular dialect should be avoided.

#### 8.3.3 Vocabulary used in prompts

Prompts should contain only vocabulary that the expected user population would be likely to understand.

EXAMPLE Jargon and technical terms are avoided unless targeted to users familiar with those terms.

#### 8.3.4 Consistency of terminology use

Terminology use should be consistent throughout an application and among applications that are integrated with each other within a system.

#### 8.3.5 Unambiguous prompts

Prompts should be stated so that they are unambiguous to the target user population(s).

#### 8.3.6 Simple linguistic forms

Simple linguistic constructions and word forms should be used whenever possible.

EXAMPLE Avoid passive voice.

#### 8.3.7 Minimum syllable count

Speech messages should contain at least four syllables to aid intelligibility.

NOTE 1 Using polysyllabic words improves speech intelligibility.

NOTE 2 Four- to eight-syllable messages present information rapidly and cause less interference with other communication than longer messages.

#### 8.3.8 Sentence usage

If possible, use short sentences rather than single words in isolation to improve speech intelligibility.

### 8.4 Voice

#### 8.4.1 Use of synthetic speech

a) Digitized rather than synthetic speech should be used.

b) If it is not possible or practical to use digitized speech, the synthetic speech used should simulate natural human speech as accurately as possible.

NOTE Synthetic speech has been found to require more attentional resources than natural speech.

### 8.4.2 Source voice identification

If the source of the speech message is relevant to the task, then that source's voice should be used if possible.

NOTE If someone the user knows is sending him or her a message, using that person's actual voice allows the user to recognize it and also interpret the message using learned cues to that person's speaking style.

### 8.4.3 Dialects and accents

Dialects and accents should not interfere with a user being able to recognize what is being said.

NOTE See 8.3.1.

### 8.4.4 Speech output rate

The speech output rate should be adjustable, with a default setting of 160 words per minute.

NOTE Listeners with visual impairment often prefer speech to be much faster than this default setting (i.e. up to 400 words per minute).

### 8.4.5 Intonation and prosody

To increase clarity, the prosody, including intonation patterns, of information items should be similar to the prosody in human speech.

## 9 Earcons

### 9.1 Meaningful earcons

Where possible earcons should make use of sounds that have very clear relationships to what they represent.

EXAMPLE The sound of hands clapping is used to represent task success or approval of some action.

### 9.2 Sound organization

If a set of sounds is used to represent a wide variety of information, a grammar should be established to organize the sounds into a semantic and syntactic hierarchy.

EXAMPLE A particular or general rhythm can be used for a certain category of information, and modulations of that rhythm or different melodies can be used to represent different specific information within that category.

### 9.3 Earcon usage

If the task requires the user to work with a large amount of dynamic data, consider using earcons to reduce the user's mental workload and reduce error recovery time.

### 9.4 Earcon construction

Earcons should be constructed in such a way that they are easily discerned and discriminated. They should be easily and clearly comprehended.



 

### 9.5 Earcon duration

The earcon's duration should be just long enough to convey the associated information effectively.

NOTE If an earcon is composed of tones, the optimal earcon length is two to four tones, which tends to be long enough for users to recognize its rhythm and/or melodic character and short enough to facilitate memorability and reduce the likelihood of users experiencing audio fatigue.

### 9.6 Consistency across earcons

Earcons that have concepts in common should be similar in at least one audio dimension, preferably a dimension that consistently represents qualities in that concept which is shared.

### 9.7 Reversing earcons

If an earcon is composed of a sequence of tones, reversing the order of the tones should not be used to construct another earcon unless the original earcon has a very clearly discernible rising or falling frequency pattern.

## 10 Coding

### 10.1 General

#### 10.1.1 Meaningfulness of codes

Meaningfulness should be built into codes however and whenever possible.

NOTE Meaningfulness is increased when clear associations exist between coded information and its intended meaning. Task performance is more accurate and rapid if codes are meaningful.

#### 10.1.2 Access to meaning of code

When the meaning of a code is not obvious to the user, then information about the meaning of the code should be easily accessible.

#### 10.1.3 Rules of code construction

a) Rules should be established for the specification of codes.

b) Rules for coding should be applied consistently and unambiguously.

#### 10.1.4 Distinctiveness of codes

Codes should be used that are perceptually distinct from each other.

#### 10.1.5 Consistent coding

Codes should be used consistently with the same meaning or the same function.

#### 10.1.6 Coding relative values

When the relative (rather than the absolute) value of a variable is important, gradual changes in intensity, frequency, duration, temporal separation, or position should be used to show the relative values of the single variable.

NOTE Timbre is not be used for coding continuous values (see 9.2.4).

### 10.1.7 Ordered coding

When variation in intensity, frequency, duration, temporal separation, or position is used for coding relative values, the assigned code values should be ordered so that the extreme values in those dimensions correspond to the extreme values of the coded variable.

EXAMPLE If an event is happening sooner than expected, then the sound intensity coding progressively indicates the upcoming occurrence (reverse camera sound in case of pull back).

### 10.1.8 Change of state

If a particular auditory dimension is used to indicate a state, then a change in that same auditory dimension should be used to indicate a change in state.

### 10.1.9 Combination codes

More dimensions should be combined and fewer levels of each dimension should be used in creating memorable and meaningful codes.

### 10.1.10 User training in codes

Users should be trained to recognize the differences between each unique combination when combining dimensions for coding.

### 10.1.11 Counting simultaneous tones

The number of simultaneously presented tones should not be used to communicate a discrete value unless the number of tones is 0 or 1.

NOTE It is too difficult with most sounds to identify the number of tones being played, and discriminability depends highly on timbre and harmonic frequencies and their relation to the fundamental frequencies. Multiple simultaneous frequencies often blend to form a chord that is interpreted more easily as a single unit, even when different timbres are used.

EXAMPLE Number of simultaneous tones can be used to code true or false, but generally users are not expected to be able to count the number of simultaneous tones if it is greater than one.

## 10.2 Dimension-specific coding

### 10.2.1 Number of frequencies for pitch coding

Four to seven levels of frequency should be used for pitch coding.

### 10.2.2 Frequency selection for pitch coding

When coding based on frequency, select frequencies spaced throughout the spectrum.

### 10.2.3 Need for redundant coding when coding based on frequency

Frequency should not be used as the only dimension of coding for a particular variable.

NOTE Performance is better when selected frequencies are highly distributed (see 10.2.2), but people vary greatly in their pitch discrimination ability.