ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

**ГОСТ Р 52873— 2007**

# СИНТЕЗАТОРЫ РЕЧИ ДЛЯ СПЕЦИАЛЬНЫХ КОМПЬЮТЕРНЫХ РАБОЧИХ МЕСТ ДЛЯ ИНВАЛИДОВ ПО ЗРЕНИЮ

## Технические требования

Издание официальное

БЗ 7—2007/213

Москва
Стандартинформ
2008

gost.e-ecolog.ru

## Предисловие

Цели и принципы стандартизации в Российской Федерации установлены Федеральным законом от 27 декабря 2002 г. № 184-ФЗ «О техническом регулировании», а правила применения национальных стандартов Российской Федерации — ГОСТ Р 1.0—2004 «Стандартизация в Российской Федерации. Основные положения»

**Сведения о стандарте**

1 РАЗРАБОТАН Негосударственным учреждением «Институт профессиональной реабилитации и подготовки персонала Общероссийской общественной организации инвалидов Всероссийского ордена Трудового Красного знамени общества слепых «Реакомп» (НУ ИПРПП ВОС «Реакомп»)

2 ВНЕСЕН Техническим комитетом по стандартизации ТК 381 «Технические средства для инвалидов»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 27 декабря 2007 г. № 551-ст

4 Настоящий стандарт разработан по заказу Федерального агентства по здравоохранению и социальному развитию в рамках федеральной целевой программы «Социальная поддержка инвалидов на 2006—2010 годы», утвержденной Постановлением Правительства Российской Федерации от 29 декабря 2005 г. № 832

5 ВВЕДЕН ВПЕРВЫЕ

*Информация об изменениях к настоящему стандарту публикуется в ежегодно издаваемом информационном указателе «Национальные стандарты», а текст изменений и поправок — в ежемесячно издаваемых информационных указателях «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ежемесячно издаваемом информационном указателе «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет*

© Стандартинформ, 2008

Настоящий стандарт не может быть полностью или частично воспроизведен, тиражирован и распространен в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

## Содержание

1 Область применения . . . . . 1
2 Нормативные ссылки . . . . . 1
3 Термины, определения и сокращения . . . . . 1
4 Технические требования . . . . . 2
4.1 Общие требования . . . . . 2
4.2 Требования к входным и выходным данным . . . . . 2
4.3 Алгоритм работы программных средств . . . . . 3
4.4 Объемно-временные характеристики программных средств . . . . . 3
4.5 Требования надежности программных средств синтезатора . . . . . 3
5 Эргономические требования . . . . . 4

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

# СИНТЕЗАТОРЫ РЕЧИ ДЛЯ СПЕЦИАЛЬНЫХ КОМПЬЮТЕРНЫХ РАБОЧИХ МЕСТ ДЛЯ ИНВАЛИДОВ ПО ЗРЕНИЮ

**Технические требования**

Speech synthesizers of special computer workplaces for invalids on sight.
Technical requirements

**Дата введения — 2009—01—01**

## 1 Область применения

Настоящий стандарт распространяется на синтезаторы речи для специальных компьютерных рабочих мест, применяемые инвалидами по зрению.

## 2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

ГОСТ Р МЭК 958—93 Интерфейс цифровой звуковой

ГОСТ Р 51645—2000 Рабочее место для инвалида по зрению типовое специальное компьютерное. Технические требования к оборудованию и производственной среде

Примечание — При пользовании настоящим стандартом целесообразно проверить действие ссылочных стандартов в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет или по ежегодно издаваемому информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по соответствующим ежемесячно издаваемым информационным указателям, опубликованным в текущем году. Если ссылочный стандарт заменен (изменен), то при пользовании настоящим стандартом следует руководствоваться заменяющим (измененным) стандартом. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, применяется в части, не затрагивающей эту ссылку.

## 3 Термины, определения и сокращения

3.1 В настоящем стандарте применены следующие термины с соответствующими определениями:

3.1.1 **программа экранного доступа:** Программное средство для чтения текстовой информации и озвучивания действий пользователя с помощью синтезаторов речи и/или отображения того же посредством брайлевских дисплеев.

3.1.2 **синтезатор речи**: Комплекс технических и программных средств, преобразующих текст, составленный на различных языках, в звуковой сигнал, воспринимаемый слушателем как аналог человеческой речи при различных степенях разборчивости и естественности звучания.

Примечание — Исходный текст передается для преобразования в цифровом виде с различных носителей информации или по цифровым линиям связи. Синтезаторы речи могут быть разделены на моноязычные (например, только русский язык) и многоязычные (например, русский и английский и др.); программные (функционирующие на базе технических и программных средств) и аппаратные (действующие на базе технических средств). Синтезаторы речи следует отличать от устройств речевой индикации, которые могут выдавать речевые сообщения, состоящие из ограниченного числа слов и фраз, начитанных диктором и хранящихся в памяти устройства (в том

**Издание официальное**

числе персонального компьютера), например в тифлоприборах (говорящие часы, тонометр, озвучивание клавиатуры лифта и т.д.).

3.1.3 **разборчивость речи:** Обеспечение правильного воспроизведения и корректного распознавания речи при прослушивании, в том числе и при повышении скорости воспроизведения (определяется экспертным методом).

3.1.4 **специальное компьютерное рабочее место инвалида по зрению:** Рабочее место, включающее в себя комплект средств вычислительной техники на основе персонального компьютера, оснащенного специальными устройствами и специальным программным обеспечением, позволяющими инвалиду по зрению выполнять профессиональные обязанности.

3.1.5 **просодическая информация:** Преобразованный в последовательность звуков текст.

3.1.6 **цифровая фильтрация:** Обработка отсчетов цифрового сигнала для получения требуемой амплитудно-частотной характеристики сигнала.

3.1.7 **дифоны:** Участки речевого сигнала, включающие в себя переходы между звуками.

3.1.8 **фонема:** Минимальная смыслоразличительная единица устного языка, имеющая временну́ю упорядоченность.

3.1.9 **аллофоны:** Фонетически обусловленные (специфические) комбинаторные и позиционные варианты данной фонемы.

3.1.10 **субаллофоны:** Составные части аллофонов, необходимые для формирования их звуковых сочетаний.

3.1.11 **формантный метод:** Метод формирования речевого сигнала, при котором звук формируется формантным вокодером (без использования базы естественных звуковых единиц) на основе поступающей к нему просодической информации.

3.1.12 **параметры синтезируемой речи:** Изменяемые характеристики синтезируемой речи (тип голоса, скорость речи, тон и тембр голоса, ударение и произношение).

3.1.13 **операционная среда:** Комплекс программных и технических средств, обеспечивающих выполнение прикладных программ.

3.2 В настоящем стандарте применены следующие сокращения и условные обозначения:

SAPI (Speech Application Programming Interface) — программный интерфейс речевых приложений, используемый в операционных системах Microsoft Windows;

ANSI — тип кодировки текстовых данных в операционных системах Microsoft Windows;

PCM — формат представления данных.

## 4 Технические требования

### 4.1 Общие требования

4.1.1 Синтезатор речи (аудиодисплей) применяют в составе специального компьютерного рабочего места инвалида по зрению по ГОСТ Р 51645.

4.1.2 Программные средства синтезатора должны соответствовать ГОСТ Р МЭК 958.

4.1.3 Качество синтезируемой речи должно быть таким, чтобы пользователь не испытывал напряжения при прослушивании речевых сообщений.

4.1.4 Интерфейс программных средств синтезатора должен быть разработан в соответствии с требованиями SAPI версии 5.1. Выполнение требований SAPI версии 5.1 должно гарантировать успешную работу программных средств синтезатора с Microsoft Windows приложениями.

### 4.2 Требования к входным и выходным данным

4.2.1 Программные средства синтезатора речи должны поддерживать формат входных текстовых данных в кодах ANSI (Win-1251). Программные средства синтезатора речи должны автоматически определять тип кодировки.

4.2.2 Формат выходных звуковых данных должен иметь следующие параметры:

- отсчеты сигнала — в формате PCM 16 разрядов, моно;
- частота дискретизации от 8000 до 44100 Гц.

Примечания

1 Рекомендуемая частота дискретизации 22000 Гц.

2 Рекомендуется наличие внутренней возможности генерировать звук на стандартных частотах 11025, 16000, 22050 Гц.

4.2.3 Программные средства синтезатора речи должны позволять:

а) выбирать тип голоса диктора (рекомендуется два мужских и два женских);

б) изменять громкость, скорость воспроизведения речи, подстраивать частоту основного тона, тембр, эффект реверберации (настраиваемые параметры области регулирования).

Примечание — Подстройка тембра должна осуществляться путем цифровой фильтрации нижних и/или верхних частот с переменным коэффициентом усиления.

**4.3 Алгоритм работы программных средств**

4.3.1 Программные средства синтезатора речи должны состоять из следующих блоков:

- текстовый процессор;
- транскриптор;
- звуковая база;
- формирователь речевого сигнала;
- блок вывода звука.

4.3.2 Текстовой процессор, который получает в качестве входной информации текстовые данные, должен выполнять следующие функции:

а) выделение предложений в тексте и разбивка их на синтагмы;

б) прочтение заголовков электронных писем и документов;

в) анализ использования «е» и «ё»;

г) выполнение задачи омонимии (замóк или зáмок);

д) расстановка ударений в словах, при этом должна быть предусмотрена возможность использования словаря ударений;

е) определение типа интонации предложения, формирование интонационного контура;

ж) определение места установки и длительность пауз в чтении.

4.3.3 Транскриптор должен преобразовывать полученный от текстового процессора нормализованный текст с ударениями, расставленными паузами, интонационным контуром в последовательность просодической информации, т.е. преобразовывать текст в последовательность звуков. Каждому звуку должны соответствовать требуемые длительность и частота основного тона.

4.3.4 Формирователь речевого сигнала на основе потока просодической информации от транскриптора, используя звуковые элементы из звуковой базы, должен сформировать речевой сигнал. Работа формирователя речевого сигнала зависит от того, какой тип звуковой базы используется и есть ли она вообще. Если звуковая база не используется, то синтезатор должен формировать звуковые элементы сам на основе формантного метода (формантный синтезатор). Такой метод синтеза самый экономный, но качество речи невысокое (сильно роботизированное).

4.3.5 Рекомендуется использовать синтез речи на основе звуковых баз, состоящих из дифонов, аллофонов, субаллофонов или смешанного типа, т.к. по качеству такой метод намного эффективнее формантного и обеспечивает хорошую разборчивость.

Примечание — Степень естественности речи очень сильно зависит от качества и состава собранных звуковых единиц, а также используемых алгоритмов модификации звука по частоте основного тона и длительности.

4.3.6 Программные средства синтезатора должны быть реализованы в виде отдельной программы и должны иметь блок вывода звука. С помощью соответствующих стандартных интерфейсов программные средства синтезатора должны формировать звуковой поток данных на звуковую карту компьютера.

**4.4 Объемно-временные характеристики программных средств**

4.4.1 Скорость преобразования текста в звук должна превышать скорость звучания. Программные средства синтезатора должны преобразовывать текст быстрее, чем он выдается через звуковую карту. При этом процент загрузки центрального процессора не должен превышать 25 % (при тактовой частоте процессора 3 ГГц).

4.4.2 Объем требуемой оперативной памяти не должен превышать 120 Мб. Размер файлов программных средств синтезатора не ограничивается. Ограничения на объем памяти и файлов программных средств синтезатора могут быть установлены для программных средств синтезатора, реализуемых на платформе с ограниченными ресурсами (Palm, карманные персональные компьютеры, мобильные телефоны). В этом случае ограничения диктуют выбор платформы.

**4.5 Требования надежности программных средств синтезатора**

4.5.1 Программные средства синтезатора должны обеспечивать корректную и устойчивую работу программ экранного доступа в процессе их взаимодействия.

4.5.2 Программные средства синтезатора должны обеспечивать устойчивость работы при воспроизведении текста, обрабатывать любые сочетания символов и не завершать обработку данных аварийно (зависание, зацикливание операционной среды).

## 5 Эргономические требования

5.1 Программные средства синтезатора речи должны включать в себя возможность регулировки скорости речи при неизменной высоте основного тона. В качестве норматива скорости синтеза речи устанавливается средняя скорость воспроизведения на уровне 90 слов в минуту.

5.2 Программные средства синтезатора должны включать в себя возможность регулировки высоты основного тона при неизменной скорости речи.

5.3 Программные средства синтезатора должны включать в себя возможность регулировки громкости звучания.

5.4 Программные средства синтезатора должны обеспечивать необходимую степень разборчивости речи.

Примечание — Степень разборчивости речи определяется экспертным методом. Речь считается разборчивой, если она понятна пользователям при средней скорости воспроизведения в диапазоне от 80 до 180 слов в минуту.

5.5 Программные средства синтезатора должны обеспечивать пользователю комфортные условия работы, которые включают в себя такие параметры как слитность воспроизведения звуков в словах, разборчивость речи, близость голоса к человеческому и другие, т.е. характеристики, которые влияют на утомляемость пользователя при прослушивании больших массивов информации.

УДК 681.327.8:006.354 ОКС 11.180.30 35.180 Э65 ОКП 40 3320

Ключевые слова: синтезатор речи, инвалиды по зрению, программа экранного доступа, параметры и характеристики

Редактор *Л.В. Коретникова*
Технический редактор *Н.С. Гришанова*
Корректор *Т.И. Кононенко*
Компьютерная верстка *И.А. Налейкиной*

Сдано в набор 03.03.2008. Подписано в печать 19.03.2008. Формат 60 × 84 1/8. Бумага офсетная. Гарнитура Ариал. Печать офсетная. Усл. печ. л. 0,93. Уч.-изд. л. 0,60. Тираж 131 экз. Зак. 272.

ФГУП «СТАНДАРТИНФОРМ», 123995 Москва, Гранатный пер., 4.
gost.e-ecolog.ru gost.e-ecolog.ru
Набрано во ФГУП «СТАНДАРТИНФОРМ» на ПЭВМ.
Отпечатано в филиале ФГУП «СТАНДАРТИНФОРМ» — тип. «Московский печатник», 105062 Москва, Лялин пер., 6.

gost.e-ecolog.ru