ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

ГОСТ Р 52873— 2017

СИНТЕЗАТОРЫ РЕЧИ СПЕЦИАЛЬНЫХ КОМПЬЮТЕРНЫХ РАБОЧИХ МЕСТ ДЛЯ ИНВАЛИДОВ ПО ЗРЕНИЮ

Технические требования

Издание официальное

Москва
Стандартинформ
2018

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАН Федеральным государственным унитарным предприятием «Российский научно-технический центр информации по стандартизации, метрологии и оценке соответствия» (ФГУП «СТАНДАРТИНФОРМ») и 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 и услуги для инвалидов и других маломобильных групп населения»

3 УТВЕРЖДЕН И ВВЕДЕН В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 31 октября 2017 г. № 1617-ст

4 ВЗАМЕН ГОСТ Р 52873—2007

5 ПЕРЕИЗДАНИЕ. Октябрь 2018 г.

*Правила применения настоящего стандарта установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящему стандарту публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящего стандарта соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «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 (*gost.e-ecolog.ru

© Стандартинформ, оформление, 2018

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 3
4.1 Общие технические требования. 3
4.2 Требования к входным и выходным данным. 3
4.3 Алгоритм работы программных средств 3
4.4 Объемно-временные характеристики программных средств 4
4.5 Требования к надежности программных средств синтезатора 4
5 Эргономические требования 4

НАЦИОНАЛЬНЫЙ СТАНДАРТ РОССИЙСКОЙ ФЕДЕРАЦИИ

СИНТЕЗАТОРЫ РЕЧИ СПЕЦИАЛЬНЫХ КОМПЬЮТЕРНЫХ РАБОЧИХ МЕСТ ДЛЯ ИНВАЛИДОВ ПО ЗРЕНИЮ

Технические требования

Speech synthesizers of special computer workplaces for the visually impaired. Technical requirements

Дата введения — 2019—01—01

1 Область применения

Настоящий стандарт распространяется на синтезаторы речи, применяемые инвалидами по зрению в качестве компонента специального компьютерного рабочего места как средства реабилитации.

2 Нормативные ссылки

В настоящем стандарте использованы нормативные ссылки на следующие стандарты:

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

ГОСТ Р МЭК 958 Интерфейс цифровой звуковой

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 ежегодному информационному указателю «Национальные стандарты», который опубликован по состоянию на 1 января текущего года, и по выпускам ежемесячного информационного указателя «Национальные стандарты» за текущий год. Если заменен ссылочный стандарт, на который дана недатированная ссылка, то рекомендуется использовать действующую версию этого стандарта с учетом всех внесенных в данную версию изменений. Если заменен ссылочный стандарт, на который дана датированная ссылка, то рекомендуется использовать версию этого стандарта с указанным выше годом утверждения (принятия). Если после утверждения настоящего стандарта в ссылочный стандарт, на который дана датированная ссылка, внесено изменение, затрагивающее положение, на которое дана ссылка, то это положение рекомендуется применять без учета данного изменения. Если ссылочный стандарт отменен без замены, то положение, в котором дана ссылка на него, рекомендуется применять 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 **вычислительная система:** Совокупность взаимосвязанных и целенаправленно взаимодействующих средств вычислительной техники.

3.1.2 **прикладное программное обеспечение:** Совокупность программных средств, выполняющих задачи, решаемые в предметной области пользователя.

3.1.3 **программа экранного доступа:** Программное средство для чтения текстовой информации и озвучивания действий пользователя с помощью синтезаторов речи.

Издание официальное

3.1.4 **синтезатор речи:** Техническое средство, преобразующее 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 числе персонального компьютера), например в тифлоприборах (говорящие часы, тонометр, озвучивание клавиатуры лифта и т. д.).

3.1.5 **общее программное обеспечение:** Комплекс программных средств, предназначенный для управления вычислительной системой и всеми ее ресурсами.

3.1.6 **разборчивость речи:** Обеспечение правильного воспроизведения и корректного распознавания речи при прослушивании, в том числе и при повышении скорости воспроизведения, определяемого экспертным путем.

3.1.7 **специальное компьютерное рабочее место инвалида по зрению:** Специальное рабочее место для инвалида по зрению, основанное на применении компьютера и оснащенное компьютерными тифлотехническими средствами универсального назначения, позволяющими инвалиду по зрению выполнять профессиональные обязанности.

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

3.1.9 **алгоритм ритмической обработки:** Последовательность действий для преобразования текстов слоговой структуры в звук в соответствии со скоростью звучания и скоростью ритма.

3.1.10 **алгоритм просодической обработки:** Последовательность действий по преобразованию текстовой информации в звуковой формат.

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

3.1.12 **дифоны:** Участки речевого сигнала, включающие переходы между звуками.

3.1.13 **фонема:** Минимальная смыслоразличительная единица устного языка, имеющая временную упорядоченность.

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

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

3.1.16 **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.

3.1.17 **параметры синтезируемой речи:** Изменяемые характеристики синтезируемой речи (тип голоса, скорость, высота голоса, ударения и произношение).

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

3.2 Сокращения

В настоящем стандарте применены следующие обозначения и условные сокращения:

- SAPI (Speech Application Programming Interface) — программный интерфейс речевых приложений, используемый в операционных системах Microsoft Windows;
- ANSI (Win-1251) — набор символов и кодировка, являющаяся стандартной 8-битной кодировкой для русских версий Microsoft Windows вплоть до 10-й версии;
- PCM — формат представления данных;
- UTF-8 (Unicode Transformation Format, 8-bit — «формат преобразования Юникода, 8-битный») — одна из общепринятых и стандартизованных кодировок текста, которая позволяет хранить символы Юникода, используя переменное количество байт (от 1 до 6);
- UTF-16LE (Unicode Transformation Format, 16-bit — «формат преобразования Юникода, 16-битный») — один из способов кодирования символов из Юникода в виде последовательности 16-битных слов.

4 Технические требования

4.1 Общие технические требования

4.1.1 Синтезатор речи, являющийся элементом аудиодисплея наряду со звуковой картой, устройством вывода звуковой информации и программой экранного доступа, применяют в составе специального компьютерного рабочего места инвалида по зрению по ГОСТ Р 51645.

4.1.2 В соответствии с ГОСТ Р МЭК 958 интерфейс программных средств синтезатора должен быть разработан под требования, предъявляемые в операционных средах, в которых он будет использован. Выполнение этих требований должно гарантировать успешную работу программных средств синтезатора с приложениями в соответствующих операционных средах, в частности: в среде MS Windows синтезатор должен иметь интерфейс SAPI версии 5.1 или выше. При этом он может поддерживать работу и с другими речевыми интерфейсами в этой системе (например, Microsoft Speech Platform).

4.2 Требования к входным и выходным данным

4.2.1 Программные средства синтезатора речи должны поддерживать формат входных текстовых данных в кодах, используемых в той операционной среде, для которой они разработаны. Наиболее распространены UTF-8 и UTF-16LE.

4.2.2 Формат выходных звуковых данных должен иметь следующие параметры:

1) отсчеты сигнала должны быть в формате PCM 16-разрядов, моно;

2) частота дискретизации — от 8000 до 44100 Гц.

Примечания

1 Рекомендуемая частота дискретизации 22 кГц.

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

4.2.3 Программные средства синтезатора речи должны позволять изменять громкость, скорость воспроизведения речи и подстраивать частоту основного тона (высота голоса).

4.3 Алгоритм работы программных средств

4.3.1 Программные средства синтезатора речи должны состоять из следующих блоков:

а) текстовый процессор;

б) транскриптор (может быть в одном блоке с текстовым процессором);

в) звуковая база (если метод синтеза использует ее);

г) формирователь речевого сигнала;

д) блок вывода звука.

4.3.2 Функции текстового процессора, который получает в качестве входной информации текстовые данные, следующие:

а) выделение предложений в тексте, разбивка предложений на синтагмы;

б) прочтение заголовков электронных писем и документов;

в) анализ использования «е» и «ё»;

г) выполнение задачи омонимии (замо'к или за'мок);

д) расстановка ударений в словах, 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 должна быть определена требуемая 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 используется тип звуковой базы и есть ли она вообще. Если звуковая база не используется, то синтезатор должен сформировать звуковые элементы на основе формантного метода (формантный синтезатор). Такой способ синтеза самый экономный, но качество речи невысокое (излишне роботизированное).

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

Примечание — Степень естественности 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 встроенную 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 и количестве ядер процессора, заявленных в минимальных требованиях к операционной среде, для которой он разрабатывается).

4.4.2 Объем требуемой оперативной памяти не должен превышать 10 % от объема, заявленного в минимальных требованиях к операционной среде, для которой он разрабатывается. Размер файлов программных средств синтезатора не ограничен. Ограничения на объем памяти и файлов программных средств синтезатора могут устанавливаться для программных средств синтезатора, реализуемых на платформе с ограниченными ресурсами (планшетные ПК, смартфоны). В этом случае ограничения диктует выбор платформы.

4.5 Требования к 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 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 голоса при неизменной скорости.

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 определяют экспертным путем. 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 др., т. е. характеристики, которые влияют на утомляемость пользователя при прослушивании больших массивов информации.

УДК 681.327.8:006.354 ОКС 11.180.30 Э65 ОКП 40 3320 ОКПД2 26.20.16.140

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

Редактор *Л.В. Коретникова*
Технический редактор *И.Е. Черепкова*
Корректор *С.И. Фирсова*
Компьютерная верстка *Е.А. Кондрашовой*

Сдано в набор 08.10.2018. Подписано в печать 23.10.2018. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 1,40. Уч.-изд. л. 1,26.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении ФГУП «СТАНДАРТИНФОРМ»
для комплектования Федерального информационного фонда стандартов,
117418 Москва, Нахимовский пр-т, д. 31, к. 2.
gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru