



Озвучивание текстов и распознавание голоса

# Компьютер, почитай мне сказку



Все чаще начинают появляться свободно распространяемые утилиты распознавания голоса, практически ни в чем не уступающие коммерческим аналогам. Правда, некоммерческие голосовые решения существовали и раньше, но только как оболочки для движков с агентами или без таковых. Реальной же бесплатной альтернативы не было.



## Полезные ссылки

Если какая-то из программ у вас не будет работать — загрузите и установите набор голосовых библиотек Microsoft:

► <http://agent.microsoft.com/agent2/sdk/lliset.exe> (6,3 Мбайт).

Поддержка русского языка:

► <http://activex.microsoft.com/activex/controls/agent2/httpsrur.exe> (3 Мбайт);

► [http://activex.microsoft.com/activex/controls/agent2/tv\\_enua.exe](http://activex.microsoft.com/activex/controls/agent2/tv_enua.exe) (1,2 Мбайт)

Что касается «агентских» программ, то для них потребуется отдельная поддержка:

► <http://activex.microsoft.com/activex/controls/agent2/msagent.exe> (400 Кбайт).

► <http://activex.microsoft.com/activex/controls/agent2/AgtX0419.exe> (128 Кбайт).

**В**озьму на себя смелость утверждать, что хотя сейчас в мире разработки ПО наблюдается относительное затишье, новые версии программ появляются достаточно регулярно, и разработчики не забывают о своих детищах.

Несмотря на то что пока так и не произошло качественного прорыва в такой интересной и перспективной области разработки, как распознавание речи и голосовое управление, новые утилиты для существующих технологий выпускаются с завидным постоянством. Среди них можно встретить и те, которые были созданы российскими разработчиками, причем они ничем не хуже зарубежных аналогов. В представленном ва-

шему вниманию обзоре я не ставил перед собой задачи выявить одного победителя — все его участники предназначены для решения разных задач, следовательно, и работают по-разному.

Каждая из этих утилит имеет определенные недостатки, и решать какие-либо серьезные задачи с помощью голосового управления невозможно. Но даже те функции, которые имеются, ясно говорят о том, в каком направлении будут идти разработки методов управления компьютером. Думаю, что от возможности командовать своим «железным другом», вальяжно расположившись в кресле, никто не откажется.

■ ■ ■ Александр Евдокимов

# GUI Reader 2.0

## ПЛЮСЫ/МИНУСЫ

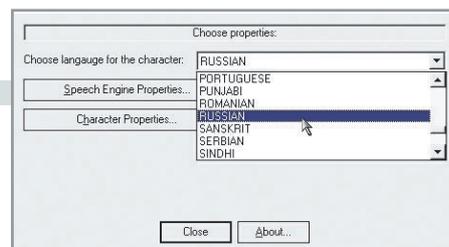
- + нетребователен к ресурсам
- + можно выбрать язык для текста
- есть только общие системные настройки

Отрадно, что и наши российские разработчики не забывают о любимом в народе голосовом ПО, регулярно выпуская интересные утилиты, причем в большинстве случаев превосходящие зарубежные аналоги. Так, скажем, GUI Reader Андрея Ефимова поддерживает все необходимые функции для подобного рода программ, но при этом занимает в пять (!) раз меньше оперативной памяти, нежели Web Talkster и CyberBuddy. 800 Кбайт против 4 Мбайт — очевидное, согласитесь, преимущество.

В состоянии покоя GUI Reader представлен иконкой в трее и никак не мешает работе пользователя. Для его ак-

тивизации нужно кликнуть правой кнопкой мыши по этому значку и выбрать в выпадающем меню то, что требуется. Можно попросить GUI Reader озвучить TXT-файл («Read File») или содержимое Буфера обмена («Read Clipboard»).

К сожалению, настроек в этой программе почти нет. Те же, что имеются, представляют собой стандартные системные меню «Свойства персонажа» и «Дополнительные параметры персонажей». Только так здесь можно настроить агента и его голос. Зато в GUI Reader разрешено выбирать язык озвучиваемого текста. Вы сами можете выставить хоть финский, хоть грузинский или суахили — главное, чтобы в вашей системе был установлен соответствующий голосовой движок с поддержкой необходимого вам языка.



Дистрибутив GUI Reader вы найдете на сайте разработчика: <http://efimov.media-pg.ru/programs/guireader/download/GUI%20Reader%20V2.00.EXE>. Оттуда можно загрузить все необходимое включая агента Джина. Нет только голосовых движков, их придется устанавливать отдельно.

## GUI Reader 2.0

Разработчик ▶ Андрей Ефимов

Сайт разработчика ▶ [www.media-pg.ru](http://www.media-pg.ru)

Размер дистрибутива ▶ 3,7 Мбайт

Условия распространения ▶ freeware

Операционная система ▶ Windows

# Web Talkster 3.9

## ПЛЮСЫ/МИНУСЫ

- + добавляет голосовые функции в Internet Explorer
- + использует анимированные персонажи
- нельзя точно настроить параметры чтения
- требует много оперативной памяти

Пожалуй, ни одна программа под Windows не имеет так же много вариаций и надстроек, как MS Internet Explorer. Что послужило причиной тому — недостатки ли самого браузера или, напротив, его огромная популярность среди пользователей, — сказать трудно, но факт остается фактом: месяца не проходит, чтобы не появился очередной клон «Исследователя».

И было бы, на мой взгляд, удивительно, если бы разработчики голосового софта остались в стороне от всеобщего увлечения программирующей молодежи и не создали бы «читающие» прибамбасы для Internet Explorer. Не скомпилировали бы, например, Web Talkster — бесплатный браузер с поддержкой технологий text-to-speech и MS Agent.

Построен он, соответственно, на движке Explorer, но занимает при этом в полтора раза больше оперативной памяти, нежели майкрософтовский прародитель.

Окно этого браузера, несмотря на объемные кнопки и возможность разукрашивания в разные цвета, все равно напоминает «Исследователя». Правда, ряд важных функций (копирования, печати, поиска на странице и т. д.) спрятан в меню «Options -> File», но все остальное сделано, как в оригинале. Плюс три новых кнопки: «Speak» (чтение всего текста на открытой web-страничке), «Selected» (чтение только выделенных абзацев) и «Quit Spk» (отмена озвучивания).

Web Talkster использует симпатичных мультипликационных персонажей — агентов. Поддержка этой технологии вместе с образом мага Мерлина устанавливается в процессе инсталляции браузера. Впрочем, библиотеки MS Speech API и английский вариант Lernout & Hauspie тоже будут предложены. Ваша задача, как и в случае со Speakopia, предварительно установить подходящий русский вариант голосового движка.

Сам же голос, точнее его тембр и скорость чтения, можно подстроить в меню «Options -> Speech». К сожалению, настройка возможна только приблизительная. Поэтому на уровне скорости «Medium» браузер Web Talkster будет читать тексты слишком быстро, а в варианте «Slow» чересчур медленно.



С помощью меню «Options -> Character Options» вы сможете также оптимизировать своего агента, выбрать персонажа («Select character») или вовсе его отключить («Move -> Off screen»). Интересно, что и без изображения агента Web Talkster будет продолжать зачитывать странички.

Из дополнительных возможностей «говорящего» браузера отмечу встроенную функцию поиска на Google.com, как в браузере Opera, а также небольшой блокнотик для «путевых» заметок «My Surf'n Notes».

Загрузить Web Talkster можно прямо с сайта разработчика: [www.code-it.com/downloads/WT\\_full.exe](http://www.code-it.com/downloads/WT_full.exe).

## Web Talkster 3.9

Разработчик ▶ Code-it Software

Сайт разработчика ▶ [www.code-it.com](http://www.code-it.com)

Размер дистрибутива ▶ 9 Мбайт

Условия распространения ▶ freeware

Операционная система ▶ Windows

# CyberBuddy 2.12

## ПЛЮСЫ/МИНУСЫ

- + многофункциональность
- + наличие встроенного почтовика и органайзера
- + гибкая система настроек
- не понимает кодировку KOI8-R

В этом году любимый мною Agent Reader ([www.aldoatools.com](http://www.aldoatools.com)) перешел в разряд shareware, без боя уступив лидерство в классе бесплатных голосовых программ виртуальному секретарю Джона Дефино CyberBuddy. И по праву. Эта утилита была весьма существенно переработана в последних версиях. Во-первых, в ее настройках появилась возможность выбора ключевых фраз для разных языков, включая русский («General -> Advanced Options -> Select Language Files»), а во-вторых, исправлена неприятная ошибка, связанная с отсутствием корректного (без нажатия «Ctrl+Alt+Del») выхода из программы по завершении работы. Теперь, избавившись от грехов молодости, CyberBuddy представляет собой мощный универсальный программный продукт, выполняющий множество полезных для широкого круга пользователей функций: «говоря-

щего» почтовика («E-mail»), органайзера («Remind»), интернет-пейджера («Buddies -> Send»), агента новостей («News»), а также наблюдателя за изменениями на web-страницах («WWW») и колебаниями курсов на биржах («Stocks»). И все это озвучено, украшено анимированными персонажами и абсолютно бесплатно. Согласитесь, есть смысл попробовать: <http://thecyberbuddy.com/cyberbuddysetup.exe>. Только вот голосовые движки и библиотеки придется закачивать отдельно: в дистрибутив CyberBuddy они не входят.

В этой программе все можно перестроить на свой лад. Так, на закладке «General -> Character and Speech Options» обязательно отрегулируйте «под себя» различные параметры агентов и голосовых движков. На странице «General -> Advanced Options» выберите скин для основного окна программы и напишите, если хотите, свои фразы для анимированных персонажей («Custom Message»). На закладке «E-mail» вам необходимо будет указать адрес своего почтового ящика, а также логин и пароль доступа к нему. Теперь по команде «Get E-mail Now» вы будете полу-



чать почту, а при выделении нужного письма еще и озвучивать его содержание («Read E-mail»). Тут, правда, есть одна заковыка. Программа CyberBuddy не только не понимает русскую почтовую кодировку KOI8-R, но и не позволяет скопировать текст писем в Буфер обмена. А это значит, что нельзя использовать внешние конвертеры. Единственный выход — попросить друзей и приятелей присылать письма в формате Windows. Или использовать транслитерацию.

## CyberBuddy 2.12

Разработчик ▶ John M. DeFino

Сайт разработчика ▶ <http://thecyberbuddy.com>

Размер дистрибутива ▶ 1,9 Мбайт

Условия распространения ▶ freeware

Операционная система ▶ Windows

# Speakonia 1.03

## ПЛЮСЫ/МИНУСЫ

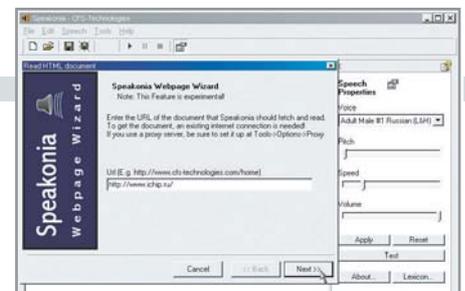
- + простой интерфейс
- + возможность записи прочитанного в WAV-файл
- настраивать можно только движки Lernout & Hauspie

Самый простой и доступный способ озвучить текст, набранный на компьютере, — воспользоваться программой Speakonia. Единственное, что от вас потребуется, — установить русифицированный голосовой движок (Lernout & Hauspie, Digalo или Sakrament). Что касается библиотек MS Speech API и поддержки американского английского языка, то они уже приложены разработчиком к инсталляционному дистрибутиву.

Интерфейс Speakonia — образец минимализма. Ничего лишнего, только основные голосовые функции. Управление точь-в-точь как в мультимедийных проигрывателях: кнопка запуска чтения, паузы и полной остановки. Параметры декламации, в принципе,

можно настраивать («Speech -> Properties»). Вы вправе выбрать движок, скорость, тембр и звучность голоса, правда, почему-то только для движков Lernout & Hauspie. Но в любом случае значения, выставленные по умолчанию, наверняка вам подойдут. Если запутаетесь с настройками, то смело кликайте по кнопке «Reset» — не ошибетесь. Тем более что качество чтения не зависит от Speakonia — подправлять можно только в тех движках, где это предусмотрено. В появившемся при щелчке мышкой по кнопке Lexicon «Редакторе пользовательских словарей» на закладке «Правка» можно уточнить неверно озвучиваемые слова или аббревиатуры. Для этого следует задать свой вариант и перевести его в специальное фонетическое письмо («Преобразовать»).

Впрочем, делать это вовсе не обязательно, поскольку Speakonia и так будет неплохо озвучивать текст из Буфера обмена, TXT-файлы, страницы в Интернете («File -> Read Webpa-



ge»), а также текущие время и дату. После регистрации на сайте программы («Tools -> Options -> Registration») появится возможность записи машинного чтения в WAV-файл («File -> Save as Wav File»).

Загрузить Speakonia можно отсюда: <http://files.cfs-technologies.com/speakonia/speakoniasetup-1.0.exe>.

## Speakonia 1.03

Разработчик ▶ CFS-Technologies

Сайт разработчика ▶ [www.cfs-technologies.com](http://www.cfs-technologies.com)

Размер дистрибутива ▶ 2,6 Мбайт

Условия распространения ▶ freeware

Операционная система ▶ Windows

# Voice Navigator 1.01

## ПЛЮСЫ/МИНУСЫ

- + высокая точность распознавания
- + можно задать голосовой ответ машины
- требует повторного обучения командам

Voice Navigator — единственная коммерческая программа в настоящем обзоре и единственная утилита, предназначенная для голосового управления компьютером. «Навигатора» создал известный петербургский Центр речевых технологий, тот самый, что занимался расшифровкой записей с затонувшей в Баренцевом море подводной лодки «Курск».

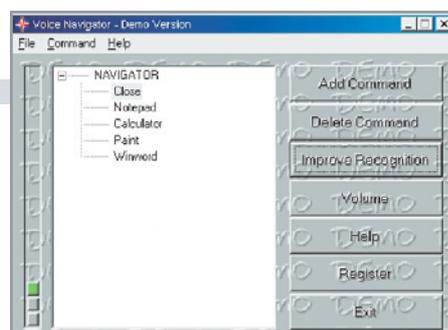
Любимое детище этого центра — «Труффальдино», комплекс голосового управления бытовой и офисной техникой. Разработчики обещают в нем полную поддержку телевизоров, телефонов и даже микроволновых печей. Правда, о том, как это выглядит на практике, я могу лишь догадываться — демоверсия «Труффальдино» ничего такого не демонстрирует. Но команды там наверняка звучат в духе известного анекдота: «Телефона, телефона, моя кушать хочет!»

А вот демоверсия Voice Navigator зарекомендовала себя с самой лучшей стороны. Правда, не сразу, а после небольшой тренировки.

Чтобы поупражняться, нужно выделить мышкой один из пяти доступных в демонстрационной версии пунктов, кликнуть по кнопке «Improve Recognition» и четко проговорить соответствующую команду. Это может быть любое слово, даже звук. «Калькулятор» я, например, вызываю по свистку.

Сразу после обучения попробуйте произнести в микрофон команду для вызова какой-нибудь программы, например Notepad, — и на экране тут же появится Блокнот. Скажите что-нибудь для Paint — и на Рабочий стол подадут стандартный графический редактор. Когда закончите работу, назовите условную команду для Close — и активное окно закроется. Приказы выполняются довольно быстро и точно, но время от времени их приходится повторно разучивать.

Кроме всего прочего, для команд пользователя можно задать определенный го-



лосовой ответ машины в так называемом «Sound File Manager». Если вас не устроит вариант, принятый по умолчанию, кликните «Replace» и запишите или продиктуйте свой собственный.

Демонстрационную версию «Навигатора», по-моему, найти нельзя нигде, кроме сайта его разработчиков: [ftp://ftp.speechpro.com/pub/demos/navigator\\_eng.exe](ftp://ftp.speechpro.com/pub/demos/navigator_eng.exe). Само собой разумеется, для работы с этой программой вам понадобится микрофон.

### Voice Navigator 1.01

Разработчик ▶ Центр речевых технологий

Сайт разработчика ▶ [www.stc.rus.net](http://www.stc.rus.net)

Размер дистрибутива ▶ 1,2 Мбайт

Условия распространения ▶ commercial demo

Операционная система ▶ Windows



## Java-реализация

### Солист Рабочего стола

Не знаю, плохо это или хорошо, но программы, основанные на технологии MS Agent, создаются в основном для потехи детишек разного возраста. Конечно, отдельные разработчики пытаются придать своим программам хоть какой-то ореол респектабельности и серьезности, но от этого их творения получаются еще комичнее. Другие же и вовсе не скрывают развлекательного характера своих творений, как, например, создатель Java-утилиты MUSH.

MUSH — это сокращение от Music Scripting Helper. Эта утилита создает скрипты для музыкальных MIDI- и KAR-файлов.

Пользоваться MUSH несложно. Нужно только выбрать агента (желательно совпадающего с основным), расставить галочки по всем пунктам, кроме «La' the words», решить, куда отправить готовый клип и, кликнув «Convert», найти подходящий для обработки MIDI- или

KAR-файл. Подойдет любая композиция, но только с английскими словами — русские тексты MUSH не поддерживает. Кстати, если кликнуть по кнопке «Find same Karaoke Files», в браузере откроется страничка с множеством таких KAR-файлов.

Когда определитесь, анимированный персонаж исполнит выбранную вами песенку. На Лучиано Паваротти это будет, конечно, мало похоже, но зато оригинально и по-своему где-то даже душевно. Если исполнение вас устроит — MUSH сохранит результат. Из трех возможных вариантов я лично предпочел самый простой и доступный — экспорт в Internet Explorer («Basic web output»).

Полученный HTML-файл будет состоять всего из одной кнопки «Play», кликнув по которой можно вызвать поющего агента. Программу MUSH ищите на сайте разработчика: [www.pedog.com/mush.zip](http://www.pedog.com/mush.zip) (132 Кбайт).

