

# ОЗВУЧЕНИЙ РЕДАКТОР ТЕКСТІВ ДЛЯ КОРИСТУВАЧІВ З ВАДАМИ ЗОРУ

Валерій Пулипенко, Микола Сажок

Міжнародний науково-навчальний центр інформаційних технологій та систем

40 просп. Академіка Глушкова, Київ 03680

Електронна пошта: [mykola@uasoiro.org.ua](mailto:mykola@uasoiro.org.ua), [speech@public.icyb.kiev.ua](mailto:speech@public.icyb.kiev.ua)

## Abstract

*Valery Pylypenko, Mykola Sazhok. Spoken Text Editor for Blind Users.* The presented application is developed to provide interaction between computer and blind users. By means of both spoken feedback and input user can edit formatted text, insert there objects, save and print the document. Benefits and problems of visual channel substitution with spoken one are discussed.

## Вступ

Персональні ЕОМ все ширше входять у життя людей, допомагаючи їм у праці та побуті, збагачуючи знаннями та досвідом. Проте існує чимала категорія людей, яким елементарне користування комп'ютером не доступне через різні фізичні вади.

Зокрема, гостро ця проблема постає перед людьми з вадами зору. Адже сучасні операційні системи, серед яких чільне місце посідає *MS Windows*, насправді зовсім не пристосовані під потреби таких користувачів, оскільки ОС орієнтовані на візуальну підтримку взаємодії з користувачем. Сама концепція вікон передбачає візуальне сприйняття цих об'єктів і операції над ними. Це також стосується і стандартних засобів введення інформації, таких як клавіатура та, особливо, мишка.

Проте сучасний комп'ютер є досить потужним засобом, в якому є можливість заміни візуального

каналу взаємодії звуковим. Це дозволяє на основі стандартної ПЕОМ розробити альтернативні засоби праці користувачів з вадами зору в ОС, подібних *Windows*. Прикладом такого засобу є звуковий редактор, призначений для набору, форматування, збереження та друкування текстів.

## 1 Структура редактора

Звуковий редактор складається з 5 основних програмних модулів, які взаємодіють з комп'ютерними засобами введення-виведення, такими як мікрофон, звукові динаміки та клавіатура (Рис. 1).

1. Текстовий редактор забезпечує редагування, форматування, пошук, збереження, відкриття та друк документів.
2. Координатор звукових подій та клавіатури здійснює синхронізацію введення-виведення звуку та натискань на клавіші.
3. Модуль розпізнавання голосових команд з мікрофону забезпечує розпізнавання та навчання окремо вимовлюваних слів.
4. Програмний модуль озвучення текстів, який синтезує мовний сигнал та виводить його до навушників або на зовнішні динаміки.
5. Акустична та лінгвістична бази даних озвучення текстів, яка включає модуль розмітки наголосів та інтерпретації іншомовних символів.

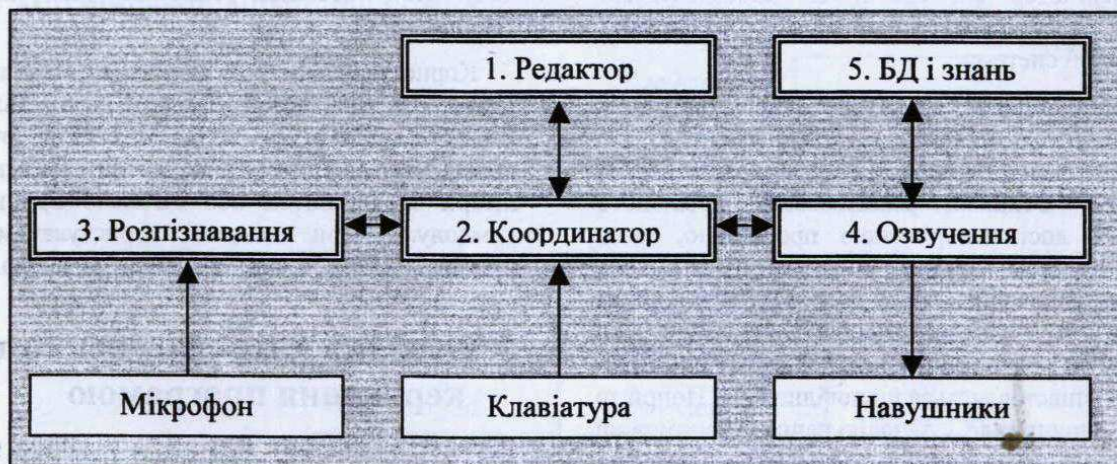


Рис. 1 Структура звукового редактора.



## 2 Засоби синтезу і розпізнавання усної мови

Синтезована усна мова виконує роль заміника візуального каналу взаємодії з користувачем. Усі дії користувача мають звуковий зворотній зв'язок, як то: вимовляння окремих літер та слів при натисканні клавіш, озвучення пунктів меню, діалогових вікон та елементів у ньому.

Користувач також повинен мати уяву про поточний стан системи: який файл він редагує, де

довільні тексти при обмеженості словника модуля розпізнавання.

Для синтезу використовується синтезатор російської мови, створений на основі синтезатора розміченого тексту, розробленого в МДУ. Цей синтезатор був доповнений модулем транскрибування української та англійської мов, оскільки тексти, якими оперують користувачі здебільшого багатомовні.

Широко використовуються короткі звуки для позначення дій користувача під час редагування. Це і

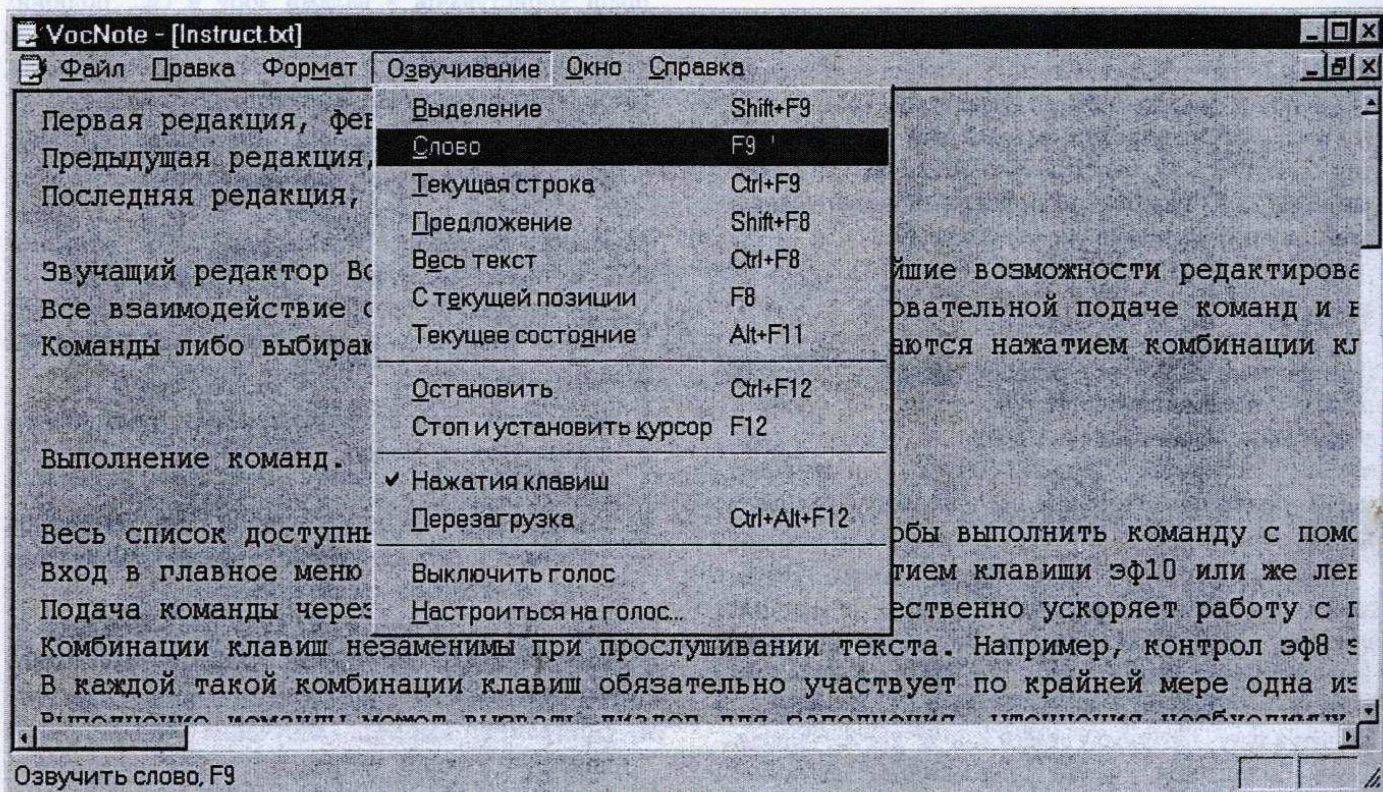


Рис. 2. Зовнішній вигляд голосового редактора користувачів з вадами зору "VocNote".

розташовується точка введення тексту та які дії очікує комп'ютер від нього. Це забезпечується введенням додаткових команд, які оголошують поточний стан системи.

Ці та інші команди користувач може активізувати як за допомогою клавіатури, так і голосом, що забезпечується використанням модуля розпізнавання усної мови. Оскільки розпізнавання довільного мовлення є досі невирішеною проблемою, то в редакторі використовується система розпізнавання окремо вимовлених слів з фіксованого словника [1]. Тому було запропоновано використовувати побуквенне надиктовування текстів, де кожній літері та символу співставляється відповідне ім'я. Попри те що це спричинило заповільнення усномовної взаємодії з користувачем в свою чергу підвищило надійність розпізнавання та дозволило вводити

зміна мови введення тексту, початок-кінець документу, особливі символи такі як "Insert", "Home", "End", "Delete" тощо.

Користувач має можливість прослухати окреме слово, речення, абзац або весь текст цілком. При переміщенні курсору відтворюється поточна буква, при цьому для синтезу використовується повна назва літери за алфавітом або загально прийнята назва символу. При введенні користувачем символу пробілу озвучується останнє введене слово.

## 3 Особливості редагування та керування програмою

Користувач з вадами зору має можливість працювати зі стандартними командами редагування, такими як набір тексту, пошук тексту, заміна, збереження



тексту на твердому дискуві, друк (Рис. 2). Також є можливість вибору різних шрифтів для редагування, в тому числі математичних. Передбачені додаткові команди для з'ясування характеристик виділених символів, як то колір, назва, висота, нахил шрифту, мову.

Дуже важливу роль виконують команди роботи з текстовими блоками. Стандартний спосіб виділення за допомогою натисненої клавіші Shift, а тим більше за допомогою мишки не підходить для користувачів з вадами зору. У редакторі реалізований додатковий набір команд для виділення і оперування блоками тексту, який виконується за допомогою натискання клавіш, вибором пунктів меню або голосом. Користувачі виявили бажання використати накопичувальний буфер обміну (кишеню), в якому вони можуть умістити декілька порцій тексту.

Операції редагування, що можуть привести до сильних змін тексту, заблоковані або запитують підтвердження на виконання. Наприклад, заблокована заміна виділеного блоку одним символом, що дозволяється у стандартних редакторах.

Всі операції з текстом: взяття в кишеню, вставка з кишені, видалення і т.п. супроводжуються короткими звуками або докладними коментарями (за бажанням користувача).

Озвучення взаємодії з користувачем також відбувається і при роботі зі вставними об'єктами, такими як редактор математичних формул.

Режими озвучення за бажанням користувача перемикаються від максимального, коли відтворюються всі символи і повідомлення редактора, до мінімального, коли всі додаткові повідомлення не озвучуються. Максимальний режим озвучення корисний для навчання редагуванню, а мінімальний необхідний для досвідчених користувачів.

## 4 Настроювання на голос користувача

Дуже гостро постає проблема навчання на диктора, оскільки система розпізнавання, використана в редакторі, є дикторозалежною.

У звуковому редакторі застосовується навчання на голос користувача за допомогою звукової підказки (Рис. 3). Користувач викликає режим навчання і програма підказує наступне слово для навчання. Після промовляння слова, система аналізує сигнал і в випадку занадто гучного, або слабкого, сигналу запрошує промовити це слово ще раз.

Але система не має можливості скасувати промову, якщо користувач промовляє зовсім інше слово ніж підказує система. Рекомендується протестувати навчання на словник, для виявлення невдалого промовляння тих чи інших слів та скоригувати навчання.

Результати настроювання автоматично зберігаються в окремому файлі та викликаються, у випадку активації голосового керування, звуковим редактором. Можливе настроювання як на одного користувача так і на декількох.

## 5 Взаємодія з середовищем Windows

Звуковий редактор реалізований у середовищі Windows та використовує стандартні засоби даної системи. Для введення та виведення звуку використовується стандартна звукова карта типу SoundBlaster. Все це дає можливість встановлювати звуковий редактор на довільний комп'ютер, який обладнаний звуковою картою, та працює в середовищі Windows.

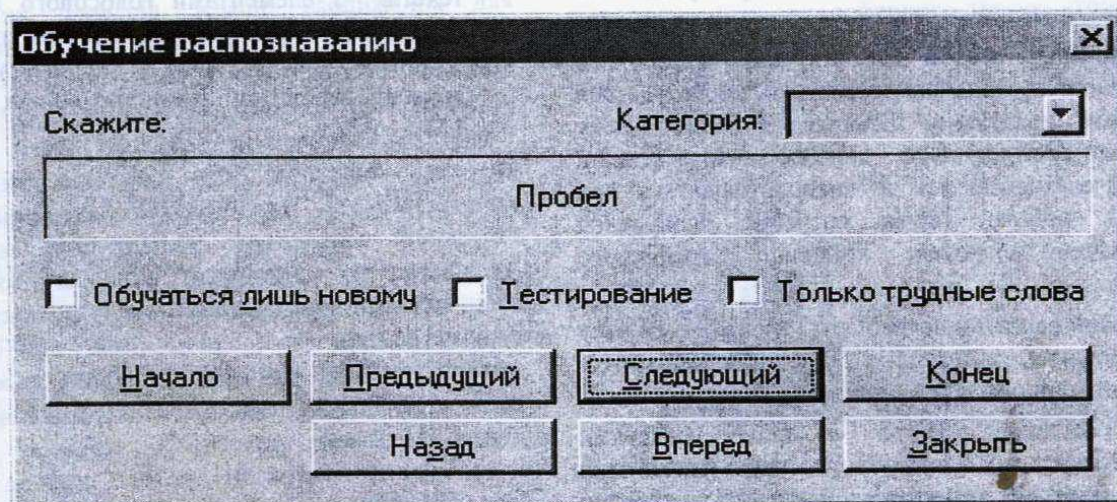


Рис. 3. Діалогове вікно для настроювання на голос користувача.



Розпізнавання та синтез мови виконуються в реальному масштабі часу та не потребують від користувача занадто довгого очікування тієї чи іншої події. Попри очевидні переваги при введенні та виведенні даних, звуковий канал не може повністю замінити візуальний за швидкістю та зручністю роботи користувача.

Віконна операційна система використовує багато візуальних засобів для підвищення працездатності та зручності користування комп'ютером. Багато з цих засобів не можливо адекватно передати в звуковому вигляді. В даному редакторі переслідується прагматична ціль - надати можливість користувачеві з вадами зору працювати з текстами.

Для повноцінної роботи користувачеві з вадами зору доводиться дуже довго опанувати концепції ОС *Windows*, прийняті для роботи з діалоговими вікнами. Тому в даному редакторі максимально збережений стиль діалогових вікон *Windows* для того, щоб користувачі поступово засвоювали цей стиль, оскільки в майбутньому вони неминуче користуватимуться ОС *Windows* і такими складними редакторами як *Word*, *Excel*.

Наприклад для відкриття файлу, його збереження використовуються стандартні діалогові вікна, в яких були прибрані всі зайві в певному розумінні елементи. Це дозволило трохи спростити процес задавання потрібних користувачеві опцій, таким чином поступово його вводячи у світ сучасних операційних систем.

## 6 Досвід користування

Даний редактор використовується членами асоціації "Вікно у світ" [2] у їх повсякденній роботі при написанні листів, статей, звітів тощо.

Користувачі мають можливість порівняти звуковий редактор з іншими засобами допомоги. Брайлевський дисплей дозволяє адекватно передавати графічні об'єкти, але має значну собівартість та малу швидкодію. Також потребує значних зусиль по навчанню користувача на брайлевській техніці.

Універсальні засоби для озвучення графічного операційного середовища дозволяють користуватися довільними додатками *Windows*. З іншого боку користувач зіштовхується з явищем занадто великої кількості елементів, які не використовуються при звичайній роботі. Також потребує опанування користувачем складних концепцій графічного інтерфейсу.

Звуковий редактор потребує значно меншого часу для освоєння, та дозволяє зробити перший крок в опануванні концепцій середовища *Windows*.

При асоціації "Вікно у світ" відкритий комп'ютерний клас, в якому користувачі навчаються роботі зі звуковим редактором. Для цього розроблені

спеціальні методики навчання, якими користуються вчителі при навчанні учнів.

## 7 Перспективи розвитку

Подальший розвиток редактора полягає у тому, щоб спростити роботу з таблицями, формулами та іншими сучасними засобами створення текстів. Такі складні об'єкти потребують розроблення додаткових засобів звукового відображення.

Надто гостро постає проблема оброблення малюнків, маніпулювання графічною інформацією, якими рясніють текстові документи.

Також необхідно створювати додаткові можливості, такі як голосове управління електронною поштою, скануванням книг і навіть навігацією в Інтернеті. Інтегрування додаткових засобів може дати можливість виконувати різноманітну корисну роботу, використовуючи єдиний засіб.

## Висновки

Реалізовано звуковий редактор текстів для користувачів з вадами зору в середовищі *Windows*. Вказана категорія користувачів комп'ютера має змогу редагувати тексти, зберігати, викликати та друкувати їх.

Візуальний канал взаємодії замінюється звуковим за допомогою синтезу та розпізнавання мови.

Звуковий редактор використовується користувачами з вадами зору у їх повсякденній роботі.

## Література

1. Винцюк Т.К. Анализ, распознавание и смысловая интерпретация речевых сигналов. — Киев: Наукова думка, 1987, 264 с.
2. Пилипенко В., Сажок М. Озвучений редактор текстів з елементами голосового керування в середовищі *Windows*. // Праці 1-го міжнародного семінару "Інформаційні технології у соціально-трудоу реабілітації інвалідів" – Київ, 2001, с 121-123.