

## Voice Windows - Voicified Script-based Win98 Platform for the Visually Disabled

TUNEYOSHI KAMAE,<sup>†1</sup> TAKAYUKI WATANABE,<sup>†3</sup> FUJIO KOIDE,<sup>†6</sup>  
TOORU KURIHARA,<sup>†2</sup> HIROHIKO HONDA,<sup>†4</sup> SHINICHIROU UNO<sup>†5</sup>  
and SAWAKO TAJIMA<sup>†7</sup>

We are developing two sets (platforms) of utility softwares with which the visually disabled can exploit Win98 as much as the sighted do by using script languages (the Voice Windows project). One platform is an extension of Windows Scripting Host (WSH) with Aural User Interfaces (AUIs) and large-format GUIs (Voice WSH), and the other, an extension of Multi-Lingual Emacs (Mule) on Win98 with AUIs (Voice Meadow). The former (Voice WSH) enables to access standard Windows98 utilities and applications through script languages. The latter (Voice Meadow) supports standard UNIX utilities including Mule, MEW, and others.

### 1. Introduction

Personal computers have long been used by visually disabled professionals and students. In the age of the Disk Operating System (DOS), I/O operations were processed by character-based interfaces. Text reading with a voice synthesizer and/or presentation on a braille display were relatively straightforward.

T. V. Raman, a visually disabled graduate student at Cornell University, developed Emacspeak<sup>\*</sup>, a speech-enabled Emacs, and opened a path for the visually disabled to become professional programmer<sup>1)</sup>. In Japan, Masao Saito who lost sight at pre-school age wrote a widely used Japanese text reader (VDM100) that has an efficient Kana-Kanji translator and a homophonic Kanji identifier<sup>2)</sup>. Usefulness of personalized computers had become widely recognized by the visually disabled through these assistive programs<sup>3)</sup>.

As Microsoft Windows95 became the dominant OS for personal computers, the situation changed. The MS-DOS on its DOS window became incompatible with many then-existing screen readers and visually impaired users began to be dissociated from the personal computer evolution. Even

the most basic application, the word processor, has been upgraded to be equipped with proportional fonts. Formatted documents now occupy a large fraction of Internet traffic, making the barrier higher for the sight disabled to participate in the information revolution.

The Voice Windows project aims at lowering these barrier by supplying a set of AUIs and large-format GUIs accessible from script-based platforms available in the standard Windows environment. They include Windows Scripting Host by Microsoft<sup>3)</sup> and Mule for Windows95/98 (Meadow) by Miyashita<sup>5)</sup>.

Windows Scripting Host (WSH) supports several script languages: including Visual Basic Script (VBScript). Meadow runs on publicly archived Emacs-Lisp source codes: their long-term stability is assured.

The present work has been inspired by the formats for audio presentation described in T. V. Raman's voice-based system, AS<sup>T</sup>ER (Audio System for TEchnical Readings)<sup>1)</sup>. According to the writing, a table should be reformatted in serial audio rendering so that correlation among data can be understood.

User-friendly aural interfaces for the WSH platform and the Meadow platform need a high quality multi-lingual text-to-speech (TTS) with multi-tasking capability. The speech system has to support Kana-Kanji translation by uniquely identifying Kanji characters, keyboard echoing and confirmation, and articulation of special characters, all concurrently with the text reading. The speech system presented here is the first that meets these requirements.

The present paper describes the overall project

†1 Department of Physics, School of Science, University of Tokyo

†2 Department of Information Sciences, Tsukuba College of Technology

†3 Information Science, Shonan Institute of Technology

†4 Institute of Space and Astronautical Sciences

†5 Faculty of Social and Information Sciences, Nihon Fukushi University

†6 Create System Development

†7 ITRC

\* [www.cs.cornell.edu/Info/People/raman/emacspeak/](http://www.cs.cornell.edu/Info/People/raman/emacspeak/).

and its basic features.

## 2. Multitasking, multimodal, and multilingual Text-to-Speech System

Microsoft's Windows98 is a multitasking OS and the TTS system are required to mix concurrent voice outputs from multiple tasks to be most effective. Such concurrency is crucial in the following cases: Key echo and system messages while listening to a text; Dialog with a Kana-Kanji translator during the text input; Sharing the audio device with background jobs. Such multitasking function has not been used in the aural interface until now.

The TTS engines are assumed to conform to the basic specifications of Microsoft Speech SDK for Win32 applications (SAPI). Several Japanese TTS engines now meet SAPI specifications at varying degree. Among them are Toshiba's Japanese TTS engine and Create System Development's TTS engine developed for the Voice Windows project. We use the English TTS engine Microsoft distributes with its SAPI as the English speech engine.

## 3. Voice Windows Scripting Host

A collection of utility programs have been written mostly in Visual Basic V.6. They have a standard graphic user interface (GUI) for users with low vision and a standard aural interface (AUI) for users without vision. Multiple choices of GUI are prepared: several background and font color combinations; and several format (font) sizes. For AUI, we have prepared 2 levels of aural guidance: the introductory level and the advanced level. Users can set the voice speed at slow (the normal speed), medium (1.5 times the normal), and fast (twice the normal). The utility programs include the Voice Windows version of basic I/O functions.

All utility programs can be called from MS-DOS and WSH. There are two ways to use them in WSH: one as in MS-DOS through the WSH shell and the other as "objects" similar to the objects in Visual Basic. These objects (functions and subroutines) have properties and methods with which programing becomes simple and straightforward.

Voice Windows include a speech-enabled script language editor: an editor that assists programing in VBScript and VBA.

## 4. Voicification of Mule for Win98: Voice Meadow

In Voice Meadow we use 'hook' and 'advice' features of Emacs Lisp to augment aural interfaces to existing Emacs Lisp functions. That is several TTS commands are added when standard Emacs Lisp functions are called. The TTS engines are the same SAPI engines used in the WSH-based platform.

At present, basic editor functions of Mule, the mail application, MEW, and some Internet utility programs have their aural interfaces.

## 5. Future Prospects

We believe the two platforms (Speech-Enabled Windows Scripting Host and Meadow) of the Voice Windows project offer an exciting possibility for the visually disabled to reach an advanced level of Win95/98 use and write application programs with standard AUIs and large-format GUIs.

For the Meadow (Emacs) platform, the asset accumulated in Emacspeak is enormous. It is most desirable that an interface be written to convert the AUI part of Emacspeak for DEC Talk for SAPI-based TTS engines on Win98 or on Linux.

**Acknowledgments** This work was supported in part by Internet Technical Research Committee (ITRC) and by Research for the Future Program of Japan Society for the Promotion of Science under the Project "Integrated Network Architecture for Advanced Multimedia Application Systems" (JSPS-RFTF97R16301).

## References

- 1) Raman, T.V.: *Audio System for Technical Readings*, PhD Thesis, Cornell University (1994); *Auditory User Interfaces -Toward the Speaking Computer-*, Kluwer Academic Publishers (1997).
- 2) Saito, M.: Software Production for the Visually Impaired, *Journal of IPSJ*, Vol. 36, No. 12, pp. 1116-1121 (1995) (in Japanese).
- 3) Windows Scripting Host by Microsoft.
- 4) Multilingual Enhancement for GNU Emacs developed by Ken'ichi Handa Mikiko Nishikimi, Satoru Tomura and Naoto Takahashi.
- 5) Meadow, Multilingual enhancement to gnu Emacs with ADvantages Over Windows, by Miyashita Hisashi and edited by Takeyori Hara.