

2ZA-4 視覚障害者のための高機能情報処理システム Voice Meadow

渡辺 隆行<sup>1</sup>、釜江常好<sup>2</sup>、小出富夫<sup>3</sup>、宇野伸一郎<sup>4</sup>、栗原亨<sup>5</sup>、本多博彦<sup>6</sup>、田島佐和子<sup>7</sup>

湘南工科大学情報工学科<sup>1</sup>、東京大学<sup>2</sup>、クリエートシステム開発(株)<sup>3</sup>、  
日本福祉大学<sup>4</sup>、筑波技術短期大学<sup>5</sup>、宇宙科学研究所<sup>6</sup>、ITRC<sup>7</sup>

1 Voice Windows

ユーザーインターフェースとして GUI (Graphical User Interface) が使われるようになってから、初心者でもコンピュータを使えるようになった。しかし、目の不自由な人には GUI は役立たない。Raman は GUI に頼らない TTS (Text-to-Speech) システム AS\_TSR[1] 及び Emacspeak[2] を開発し、AUI (Auditory User Interface) を活用した音声出力のみで複雑な情報を処理できるようにした。しかし Emacspeak には、日本語を扱えない、DECtalk という Speech Server に特化している、UNIX を主眼にしている、などの欠点がある。

そこで我々は、視覚障害者が Windows パソコンを活用できる TTS システム Voice Windows[3] を開発している。このシステムは、Microsoft の Speech API 対応、マルチタスク対応、多言語対応、audio formatting (声の種類や3次元音声出力などで音を修飾する) 活用、オープンソース、といった特徴を持ち、日英言語の電子メールの読み書きや論文執筆などの高度な情報処理を目標にしている。

マルチタスクに対応するためには、複数プロセスの音声出力を同時に処理しなければならない。このシステムでは、複数の音声 (wave stream) を同時に出力するために、WDM 版のオーディオドライバ及び DirectSound の2つの仕組みを用いた。WDM 版のオーディオドライバはカーネルモードで音を合成するので、日本語 IME の音声化が得意な TTS システム、キーボード読み上げ専用の TTS システム、アプリケーションに特化した TTS システム、などの複数のシステムを同時に立ち上げて役割分担させることもできる [4]。

図 1 に示すように、Voice Windows は TTS server といくつかの TTS client で構成される。TTS client には、Emacspeak に代わる TTS システム Voice Meadow、一般のアプリケーションを音声化した client、日本語変換時の詳細読みを担う client、キーボードを読み上げる client などがある。また Voice Windows には、音声化のための基本的な部品 (Dialog Box など) やユーティリティーなどがある。TTS server は、複数の TTS client の出力を処理できる。独立して動作する場合は、TTS client 毎に TTS server が異なってもよい [4]。

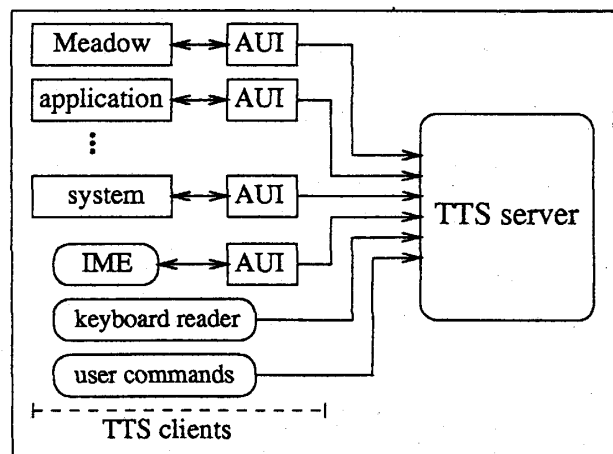


図 1: Voice Windows の構成

Voice Meadow, a high-capability information processing system for the visually disabled

T. Watanabe<sup>†</sup>, T. Kamae, T. Koide, S. Uno, T. Kurihara, H. Honda, S. Tajima

<sup>†</sup>Department of Information Science, Shonan Institute of Technology  
1-1-25 Tshujido-NishiKaigan, Fujisawa, Kanagawa 251-8511 JAPAN.  
E-mail: takayuki@la.shonan-it.ac.jp

## 2 Voice Meadow

Meadow は Windows 版の多言語対応 Emacs である。Emacs は、電子メール、Web ブラウジング、辞書検索などの多様な機能を持ち、単なるエディタを超えた優れた GNU ソフトである。新しく開発した TTS サーバー [4] を用いて Meadow を音声化するシステムが Voice Meadow [5] であり、これを用いれば、視覚障害者でも Meadow の優れた機能を利用できる。

音声化用の AUI は、上記の各機能用に Emacs Lisp で書く。このプログラムは、音声化すべきテキスト情報を Meadow から取り出し、この情報に audio formatting などの指示を追加してから TTS client (Meadow の子プロセス) の標準入力に送る。TTS client は、標準入力をプロセス間通信によって TTS server に送る。TTS server は送られた指示に基づいて、適切な声色で、同期または非同期に、送られたテキストを読み上げる。

Voice Meadow の最初のパッケージとして、電子メールソフト Mew を音声化 (Voice Mew) している。Voice Mew は Emacs Lisp の多様な関数を用いてバッファを処理し、重要なヘッダーだけを読み上げる、ヘッダーによって声や日英モードを変える、耳障りな連続文字を省略してから読み上げる、引用部の声を変える、などが簡単にできる。

## 3 まとめ

このように Voice Meadow は、晴眼者が使用している高機能なアプリケーション Meadow に AUI をかぶせて音声化している。たとえば Voice Mew の場合、専用の AUI を用意するだけで、Mew が持っているさまざまな機能を視覚障害者も使えるようになる。また Mew がバージョンアップした場合も速やかに対応できる。このようにして、視力のあるなしに関わらず、コンピュータを活用した情報処理ができるようになる。本システムはマルチタスクに対応したので、他の TTS システムとの共存や、OS のマルチタスク環境を享受できるようになった。本システムでは、日本語は日本語の TTS エンジン、英語は英語の TTS エンジンが音声化する。日英両言語をどう切り替えると聞きやすいのかなどは、検討課題である。本システムは、ソースコード付きで ITRC (インターネット技術研究委員会) 及び湘南工科大学渡辺研究室の Web サイト<sup>1</sup> で公開する予定である。

## 参考文献

- [1] T.V. Raman: "Audio System for Technical Readings", Cornell University (1994).
- [2] T.V. Raman: "Auditory User Interfaces - Toward the Speaking Computer-", Kluwer Academic Publishers (1997).
- [3] T. Watanabe, T. Kamae, T. Kurihara; "Multilingual Text-to-Speech system with active Auditory User Interface under Microsoft(R) Windows(R)", Interaction'99, (Tokyo Institute of Technology, March 4-5, 1999).
- [4] T. Watanabe, T. Kamae, T. Koide, H. Honda, S. Uno, T. Kurihara, S. Tajima: "Multilingual, Multimodal, and Multilingual Text-to-Speech System", submitted to Trans. of IPSJ (1999).
- [5] T. Watanabe, T. Kamae, T. Koide, H. Honda, S. Uno, T. Kurihara, S. Tajima; "Voicification of Meadow (Emacs), - high capability TTS server and Voice Mew -", 5th ITRC (Internet Technical Research Committee) meeting, (Sapporo, Japan, June 14-16, 1999).

<sup>1</sup><http://www.itrc.net/voicewindows/> 及び [http://www.shonan-it.ac.jp/each\\_science/info/nabeken/](http://www.shonan-it.ac.jp/each_science/info/nabeken/)