

音声入出力を利用した 視覚障害者向けWebReaderの研究

浅川智恵子・北村浩三(東京基礎研究所), 伊藤隆(ES事業)

日本アイ・ビー・エム株式会社

E-mail: chie@trl.ibm.co.jp

Web上には現在膨大な情報が存在し、その内容は多種多様にわたっている。Webには画像情報が多く含まれるが、視覚障害者にとって有効と思われる文字情報も多く存在する。これらの情報への非視覚的アクセスが実現すれば、視覚障害者にこれまでとは異なるまったく新しい情報源を提供できると期待される。しかし、現在はWindows 95への非視覚的アクセスが困難であるため、視覚障害者のWeb利用環境も整備されていない。そこで、今回ホームページ読み上げソフトであるWebReaderを試作した。

ここでは、初心者ユーザの利用を考慮し、入力にはマウスやキーボードを使用せず音声入力または数値キーパッドを利用していること、およびハイパーリンクと通常のテキストを読み分けるために男女性音での読み分けを実現した本システムのユーザー・インターフェイスについて述べる。

A Study of Nonvisual Web Access Using Voice Input and Output

Chieko ASAKAWA Kozo KITAMURA (Tokyo Research Laboratory),

Takashi ITOH (Embedded Systems Business Unit)

IBM Japan, Ltd.

E-mail: chie@trl.ibm.co.jp

Although the Web contains huge amounts of various kinds of data, it is very hard for the blind to access them, since there is no adequate environment, and few screen readers are available with Windows 95. To open up the Web as an information resource for the blind, we have built a prototype system called WebReader for reading out Web pages in synthetic speech. Among the system's many individual features are the use of voice input and a numeric keypad instead of a mouse and keyboard, and the use of a female voice for hyperlinks and a male voice for plain text.

1 はじめに

Web上には、現在多種多様にわたる情報が存在し、日々更新されている。検索エンジンなどを利用すれば、膨大な情報の中から必要な情報を容易に入手できる。Webが登場した当初、これらの情報はコンピュータ技術者にとってのみ有効と考えられていたが、今では年齢・職種を問わず、多くの人々にとって興味深い内容となっている。Webは暗眼者にまったく新しい情報源を提供したといえるが、同時に視覚障害者と暗眼者との情報格差をますます広げる結果となった。

視覚障害者は、しばしば情報障害者と呼ばれることがあるが、その言葉の通り墨字（点字と比較し、他の一般の文字をすべて墨字と呼ぶ）情報へのアクセスは容易ではない。必要な情報はすべてボランティアに依頼し、点字や録音といった媒体に変換してもらわなければならない。そのため、数日から数ヶ月待たなければならない。情報量の不足は社会参加していく上で大きな障害となっている。

これらの障害を克服するため、パーソナル・コンピュータ（以下、パソコンと呼ぶ）の普及と共に様々な視覚障害者向け機器が開発され、製品化されてきた。[1,2,3]

その結果、DOSの環境下で稼動するスクリーン・リーダーを使って、画面に表示されている文字情報を音声や点字で読むことができるようになった。CD-ROMやパソコン通信などを介して、様々な墨字情報に独力でアクセスできるのである。しかし、これらの操作法を取得するためには、パソコンの基礎知識が必要で、ユーザー層の広がりには限界があった。また、Windows 95の普及により、DOSのソフトウェアの販売停止や、DOSマシンの入手が困難になってきたなどの問題があり、視覚障害者の墨字へのアクセシビリティを向上するためにパソコンが十分にその役割を果たしているとは言えないのが現状である。

そこで、Windows 95への非視覚的アクセスの実現が課題となってきた。日本では、95Reader[4]と呼ばれるスクリーン・リーダーがすでに製品

化されており、他にもいくつかの類似システムが研究開発されている[5]。筆者らは、インターネット/Webへの非視覚的アクセス実現のための研究を進めることにし、Windows 95環境下における基本的な操作の音声サポートについては、市販されている95Readerを利用することにした。

Webには画像情報が多く含まれているが、同様に視覚障害者にとって有効と思われる文字情報も存在する。この情報への容易なアクセスが実現されれば、視覚障害者にWebというこれまでとはまったく異なる新しい情報源を提供できると期待される。この期待が実現可能なものであるかを実験・証明する目的で、今回WebReaderの試作を試みた。本システムはNetscape Navigator（以下Netscapeと呼ぶ）のPlug-Inを利用して開発されており、現在アクセス中のHTMLファイルを音声で読み上げる。他の類似ソフトウェアと異なる点は、視覚障害者の利用を考慮し、入力にはマウスを一切使用せず音声コマンドや数値キーパッドを利用すること、および画面を見ることなくその内容を読み上げたりリンクをたどることができる点にある。

本稿では、Webへの非視覚的アクセスの現状と問題点を分析した後に、システムの概要を述べる。次に実験および評価結果を報告し、今後の課題をまとめる。

本研究を通して、課題は残されているものの、WebReaderを使用すれば、コンピュータの基礎知識がない視覚障害者ユーザでも容易にWebにアクセスし、情報を入手できるようになることが明らかになった。

2 非視覚的Webアクセスの現状と問題点

すでに存在するDOS用のスクリーン・リーダーを用いて、Webにアクセスする試みが実施されている。1つはDOSからUNIXマシンにTelnetでログインし、UNIXのコマンドラインから「Lynx」というテキストブラウザを使用する方法である。Lynxを使用すれば、上下左右の

カーソル移動キーを押すだけでハイパーリンクへのジャンプやそこで指定されているリンク先のページを呼び出すことができる。このLynxの基本的な操作方法は、Webへの非視覚的アクセスを実現する上で参考となるインターフェイスである。しかし、DOSからUNIXマシンにTelnetでログインし、UNIXのコマンドを操作してネットサーフィンを行なうことは、決して一般的な利用法とは言えない。DOS上で稼動するLynxも存在するが、DOS環境下でのネットワークの設定は非常に複雑である。多くの視覚障害者の新しい情報源としてWebを位置づけるためには、そのセットアップおよびアクセス方法は、晴眼者の場合と同様あるいはそれ以上に容易なものではなければならない。

もう1つの方法は、パソコン通信が提供するテキストブラウザのサービスを利用することである。この場合のアクセス方法は、Lynxの場合と異なりやや複雑である。一ホームページのすべてのハイパーリンクには連番が振られており、ユーザはその番号を入力することで、次のリンク先を呼び出すことができる。この方法では各番号に対応するハイパーリンクを常に確認しながらネットサーフィンを行なわなければならないため、あまり実用的とは言えない。

上記いずれの場合も、読み上げ可能な範囲は画面に表示されている内容だけである。DOSのテキストモードで表示可能な文字数は、最大80文字×24行である。したがってサイズの大きいホームページを読む場合は、何度も次ページ表示キーを押さなければならないので効率が悪くなる。また、テキストとハイパーリンクの読み分けもできない。さらに、Plug-Inも利用できないため、アクセス困難なページが多く存在する。したがって、現状の方法におけるWebへの非視覚的アクセスは、本来のWebの利点を活かしたものとは言えない。

一般の晴眼者ユーザは、Windows 95プリンストールマシンとインターネット・プロバイダーへのログインIDを取得するだけで容易にWebにアクセスできる。モデムやサウンドカードなどは、プラグ・アンド・プレイにより自動

設定される。ユーザが設定する必要があるのは、ネットワーク関連の情報のみである。

このような環境を視覚障害者にも同様に提供するためには、視覚障害者向けインターネットアクセスツールもWindows 95上で稼動するものでなければならない。また、一般に提供されているPlug-Inの利用を考慮し、専用のWebブラウザの開発を行なうのではなく、既存のブラウザであるNetscapeのPlug-In機能を利用してシステムを試作することが望ましいと考えた。

WebReaderでは、Netscapeの表示画面の内容だけではなく、現在アクセス中のHTMLファイルのすべてを読み上げることができる。同様に、ハイパーリンクの検索もファイルの先頭から最後まで通して行なうことができる。また、HTMLのタグを解析しているので、音声出力のみによるWebアクセスを考慮した読み上げ機能を提供できる。この方法により、現状の問題点の多くは解決できると考えられる。

3 システムの概要

本研究の目的は、視覚障害者がWindows 95の環境でWebにアクセスし、独力で必要な情報を入手できるかどうか実験するためのプロトタイプシステムを試作することにある。冒頭でも述べたように、Web上には年齢や職種を問わず多くの人々にとって有効な情報が多く含まれている。そこで、本システムのユーザー・インターフェイスは、パソコン経験の無い視覚障害者ユーザでも容易にWebにアクセスできるようになることを目標に設計された。システムの内部構造については、今後さらに必要な機能を拡張していく過程で変更される可能性があるため、本章では非視覚的Webアクセスを実現するためのユーザー・インターフェイスについて詳細に述べる。

3.1 動作環境

本システムの動作環境は、図1のようになる。

点線で示されているシステムの使用については、ユーザが自由に選択できる。

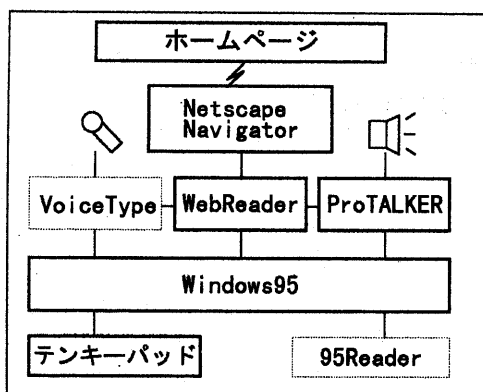


図1：WebReaderの動作環境

ハードウェアとしては、IBM PC、Aptiva、ThinkPadのPentium 100MHzプロセッサ、メモリー32MB以上での稼働を確認した。ただし、VoiceType[6]を選択する場合は、Pentium 133MHzプロセッサ、メモリー40MBで使用した。

3.2 システム起動前の操作

Windows 95起動後のインターネット・プロバイダーへの電話接続および本システムの起動は、市販されているVoiceTypeまたは95Readerの機能を利用する。95Readerのみを利用した場合、ウィンドウ間の移動やウィンドウのオープン/クローズ、スタートメニューへの移動などはすべてキーボード操作で行なわなければならない。キーボードに不慣れたユーザは、VoiceTypeのナビゲーション機能を利用すると便利である。スタートメニューへの移動は「スタートボタン」、ウィンドウ間の移動には「次のプログラム/前のプログラム」、Windowsの終了には「Windowsの終了」などといった標準的なコマンドが準備されているので、ユーザはマイクに向かって指示するだけで良い。このとき、95Readerが移動先のウィンドウタイトルを読み上げるので、併せて利用すると効果的である。

3.3 WebReaderの機能

視覚障害者がマウスを操作することは困難である。また、キーボードに不慣れたユーザにとって、キーボードを操作することは大きな負担となる。そこで、入力にはマウスやキーボードを

使用せず、音声入力または数値キーパッドを使用することとした。音声入力にはVoiceTypeのナビゲーション機能、音声出力にはProTALKER97[7]を使用している。

晴眼者との共同作業を考慮し、Netscapeの表示画面に現在アクセス中のホームページを表示する。本システムで提供しているキー操作およびナビゲーション機能を表1に示す。

表1：WebReaderのコマンドリスト

キー操作	ナビゲーション機能
0	読み始める/停止する
1	先頭へ行く
2	飛んで行く
3	最後へ行く
4	前の文
5	繰り返す
6	次の文
7	読み方
8	前のページ
9	取り消し
Enter	ストップ
.	ブックマークへ戻る
-	説明する
*	追加する
/	削除する

「0」を押すと、アクセス中のHTMLファイルの先頭から最後までを読み上げるが、この時通常のテキストは男性音、ハイパーリンクは女性音で読み上げる。これはWebReaderの特徴の一つでもあるが、ユーザはページの内容を読みながら同時にハイパーリンクも確認できる。興味のあるハイパーリンクが読み上げられた時、「2」を押すと次のリンク先を呼び出すことができるので、音声だけでも効率良くネットサーフィンを楽しむことができる。このほか、ハイパーリンクや文単位での移動、読み上げ中のポーズ機能やリスタート機能、ヒストリーをたどるなど、ネットサーフィンを楽しむ上で必要な機能は数値キーパッドに割り当てられている。また、アクセス中のHTMLファイルの先頭から最後までNetscapeの表示とは無関係に読み上げることができるので、画面表示されている範囲を考慮する必要はない。

ホームページ上に表示されている画像については、そのタグ内に<ALT>タグが記述さ

れている場合は、その文字列を読み上げる。
また、さらにその画像がハイパーリンクであった場合は女性音で読み上げる。<ALT>タグの記述がない場合は、「イメージ」またはそこで指定されているハイパーリンクアドレスを読み上げる。

HTMLファイルでは、<TABLE>タグが頻繁に使用されている。そこで、テーブルの始まりと終わりを示すメッセージを読み上げるようにし、各項目を読み上げる時には「1の1（1行1列の意）」…「2の1（2行1列の意）」などと読み上げるようにした。しかし、現実には<TABLE>タグを使用しているにもかかわらず表を表わしていない場合も多く見られる。単に画像をレイアウトする目的でこのタグを使用している場合には、テーブルを示すメッセージの読み上げはかえって理解の妨げになる。そこで、各テーブルが本来の表を表わしているかどうかシステムで判断する必要がある。現在は、表の行×列の組み合わせがn×n以下（nは変更可能）の場合は表としてみなさないよう処理できるようにした。しかし、これでは十分とは言えず、表の認識機能を強化する必要があると考えている。

表1に示しているが、数値キーパッドに割り当てられている機能はすべてナビゲーション機能でも代用できる。ナビゲーション機能を使用すると、認識に多少の時間を要するためパフォーマンスは落ちるが、マイクに向かって指示するだけでネットサーフィンを楽しむことができる。頻繁にアクセスするホームページをあらかじめ登録しておく、たとえば「日本アイ・ピー・エム」と指示するだけでそのページを呼び出すこともできる。

このほか、音声出力時の読み上げ速度、ピッチ、ボリュームも自由に設定できる。

4 実験結果および評価

本システムの有効性を確認するため、様々なジャンルからいくつかのホームページをアトラランダムに選択し、非視覚的Webアクセスの実験を行なった。

この実験は、筆者が視覚障害者であることから操作は筆者が行ない、その内容を正しく理解

できているかの確認は同僚が行なった。評価基準は、

- (1) Netscapeの画面に表示されている内容が正しく音声で読み上げられていたか
- (2) ハイパーリンクをすべて読み上げることができたか
- (3) 指定されているリンク先を正しく呼び出すことができたか

などとした。これは、前述のテーブルや画像などのGUIオブジェクトが、どの程度非視覚的Web理解に影響を及ぼしているか確認するためである。したがって、今回の実験では、VoiceTypeやProTALKERの誤認識・誤変換は評価の対象外とした。表2に実験結果を示す。

表2：WebReaderによるホームページ読み上げ精度(%)

ジャンル	表示内容の読上げ	リンクの読上げ	リンク先の呼び出し
新聞	90	90	85
教育	100	100	100
公共機関	100	100	66
テレビ	76	96	100
エンターテイメント	90	100	100
ショッピング	95	76	50
スポーツ	90	90	100
コンピュータ	86	76	76
ファッションと食	100	100	100
音楽	96	100	100
平均	92.3	92.8	87.7

ここで明らかなように、ホームページの読み上げ精度は非常に高いと言える。これは文字情報を理解する上で画像情報が大きな妨げにはなっていないことを示している。精度を下げる要因となっているのは、JavaScript、クリッカブルマップ、入力フォーム、テーブル、画像のみで構成されているページなどである。JavaScriptやサーバーサイド・クリッカブルマップについては対応できないので、晴眼者の助けを借りる必要がある。

テーブルや入力フォームについては現在最適な伝達方法を研究中である。

この実験を通して、本システムを使えば容易に多くの情報は入手できることが分かった。そこで、パソコン経験のない視覚障害者5名およびパソコン経験のある視覚障害者5名に実際に

WebReaderの被験者となってもらい、実験を行なった(表3参照)。

表3: WebReaderの評価結果

被験者	評価結果
初心者	<ul style="list-style-type: none"> ・音声入力で操作できるので簡単である ・慣れれば数値キーパッドだけでも簡単に使える ・Webとは何か少し分かったような気がする ・タイムリーな情報やあまり聞いたことのない情報を独力で読めるので楽しい ・もっと使ってみたい ・もっとインターネットやパソコンの勉強をしてみたい
経験者	<ul style="list-style-type: none"> ・テキストブラウザとは異なり気軽にネットサーフィンが楽しめる ・ハイパーリンクの区別が容易にできる ・キーボードを使っても良いのでフォームなどへの入力をしたい ・音声入力は面白いがパフォーマンスが落ちるので数値キーパッドの方が使いやすい

上記で報告した実験および評価結果から、WebReaderを使用すれば、パソコン経験の有無に関わらず視覚障害者ユーザでも容易にWebにアクセスし、情報を入手できることが分かった。

現在は実用化を目指し機能拡張しているが、多機能化に対応するため数値キーパッドへの機能割付の再設計も行なっている。また、音声出力を利用した2次元情報の伝達法についても研究中である。本システムは、各タグに対応してその伝達法を決定できるよう設計されているので、研究成果の実現と共にシステムに組み込む予定である。

5 おわりに

視覚障害者にとって、これまでのWeb上の情報はパソコン通信で得られる情報より価値があるものとは言えなかった。それは、GUI環境下への非視覚的アクセスが困難であったため、Webの利点を最大限に活かしたアクセス方法が提供されていなかったためである。このような問題は、晴眼者と同様にWindows 95環境下で一般のWebブラウザを使用し、視覚障害者に特化したWebアクセスのためのユーザー・インターフェイスが提供されれば解決できると期待

された。そこで、筆者らは視覚障害者専用のホームページ読み上げソフトウェアであるWebReaderをNetscapeのPlug-In機能を利用して試作した。WebReaderの操作では、マウスを使用せず、音声入力と数値キーパッドのみを使用した。これは、パソコン経験の無いユーザでも本システムを利用できるよう考慮したためである。また、音声出力に際してはHTMLタグを解析し、最適な情報を伝達できるようにした。これにより、パソコン経験の有無に関わらず広い範囲のユーザに容易に操作できることを確認した。

今後は、音声だけではなく点字出力も同時に提供する予定である。また、英語を始めとする外国語にも対応できるよう音声合成用エンジンの指定も可能にしたいと考えている。さらに、視覚障害者専用検索エンジンやホームページなどを準備し、WebReaderで読み上げ可能なホームページの範囲を広げ、一日も早い実用化を目指していきたい。

参考文献

- [1] 浅川智恵子, "点字英和辞書検索システム", 情報処理学会論文誌第34巻第8号, pp.1844-1852, 東京, 1993.
- [2] 浅川智恵子, "視覚障害者のためのDOS/V画面アクセスシステム", 第51回情報処理学会全国大会論文集4T-05, pp.367-368, 東京, 1995.
- [3] 斉藤正夫, "視覚障害者支援ソフトウェアの製作", 情報処理, 36, 12, pp.1116-1121, 1995.
- [4] 渡辺哲也他, "視覚障害者のパソコン利用を支援する音声合成", 信学技報 SP97-3, pp.13-20, 1997.
- [5] 岡田伸一他, "視覚障害者用スクリーンリーダーCounterVision/SR", 情処研報 95-HI-62, 東京, 1995.
- [6] 西村雅史他, "単語を認識単位とした日本語ディクテーションシステム", 音声言語情報処理 15-5, pp.27-34, 静岡, 1996.
- [7] T. Saito et al., "ProTALKER: a Japanese Text-to-Speech System for Personal Computers," IBM Research Report RT0110, June 1995.