

音声入力テキストエディタにおける指示機構とそのインターフェース

4E-3

浅貝修一朗, 原 良, 長井 純, 守屋慎次

(東京電機大学 工学部 情報通信工学科)

1 はじめに

本論文は音声入力をもちいて画面上の対象を指示(pointing)する機構を報告する。考案した指示機構はテキストエディタ上で実現した。テキストエディタにおける指示の基本機能は、利用者が望む一字を画面上で同定することである。

音声入力は、情報が1次元的に入力されるため、画面上のような2~3次元空間上の対象の指示に整合しているとは言い難い。しかし、人間によるコンピュータ操作が pointing (指示) と manipulation (操作) の繰り返しからなるため、音声入力においても指示という段階を避けて通ることはできない。したがって、1次元的な情報入力によって2~3次元空間上の対象を指示可能な、できる限り不整合を意識させな

い機構を考案する必要がある。このような問題を提起し解決案を示したのは筆者ら(文献)が最初である。本システムが対象とする文章は論文など章立てされたテキストファイルである。

2 音声入力テキストエディタ

Fig.1は試作した音声入力テキストエディタ画面例である。①はテキストデータを表示している。②は表示されているテキストデータの相対行番号である。③はスクロールである。④は、ファイルの「新規作成」「開く」「保存」である。⑤のラジオボタンにより、カーソルの単位を選択できる。⑥のラジオボタンを用いて、「索引」したい単語が選択できる(索引機能は次章で述べる)。⑦はカーソルを自動または手動で

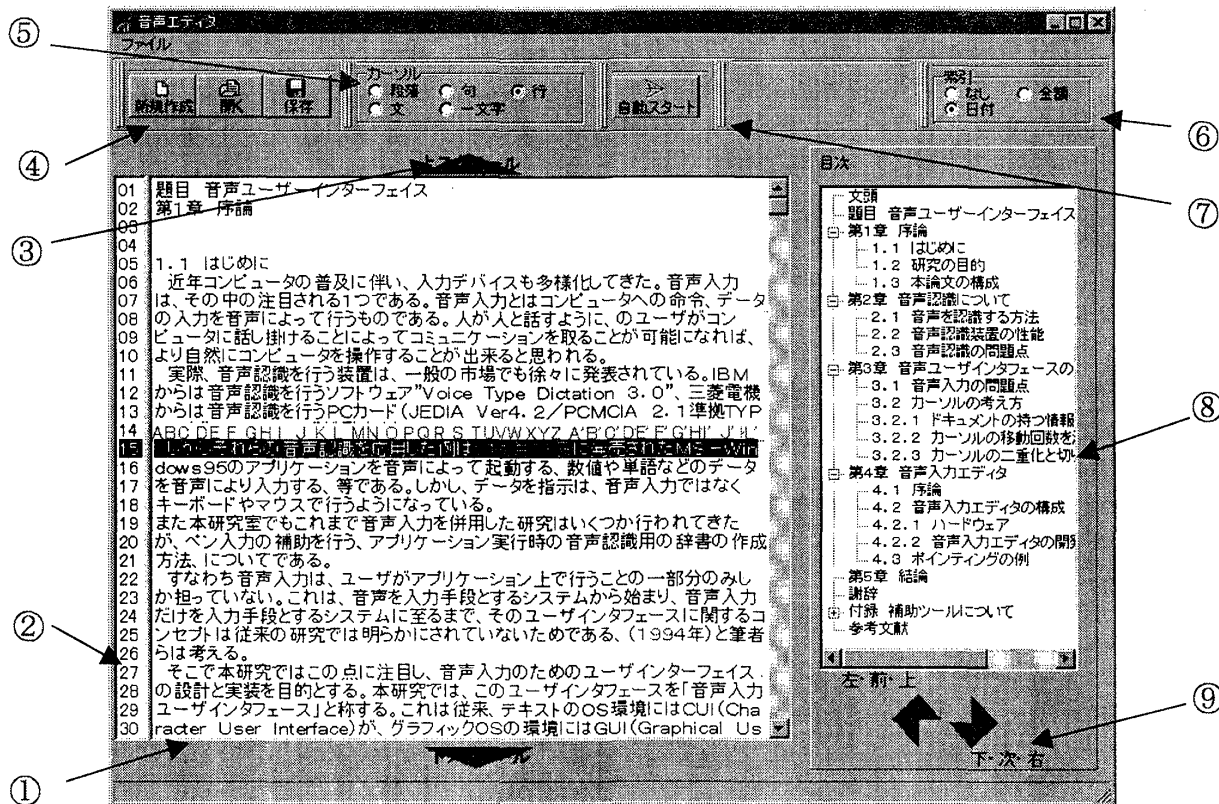


Fig.1 音声入力エディタの実行画面例



Fig.2 索引が日付の場合の行カーソル例(Fig.1の拡大図)

移動するモード選択用。⑧にはテキストデータ中の目次(章節の階層構造)が自動表示される。⑨はカーソル移動の制御指令である。

3 音声アクセス機構

音声による1次元的情報入力により、画面の2次元テキストデータを、より効果的に指示するためのさまざまな機構(音声アクセス機構)を考案した。機構は「音声カーソル」「目次」「索引」の三つに分類される。

「音声カーソル」は、通常のエディタにおけるカーソルの単位を豊富にし、単位間の階層を利用できるようにしたものである。通常カーソルは字単位であるが、音声カーソルでは段・文・句・字・行の単位で移動できるようにした。

単位間の階層については、段・文・句・字・行単位における親子関係を、指示に利用している(Fig.2)。Fig.2は、Fig.1中で反転表示している横一行(行カーソル)とその上の一行を拡大したものである。上側のアルファベット行は、下側の文章行の文字位置をしめす一種のアドレス(音声アドレス)である。Fig.1の行番号「15」を発声し、Fig.2のアルファベット「I」を発声すると文字「音」に字カーソルが移動する。この例では、行カーソルが表示される時は常に、その子単位である字アドレスが表示され、より少ない発声で字を同定できるようになっている。段カーソルのときは、その段内の任意の文が同定できる。文カーソルのときは、その文内の任意の句が同定できる。以下同様である。このように子アドレスは、目的地をカーソル内に捉えた後に効率よく指示するための機構である。目的地をウィンドウ内に捉えた後、最小二回の発声だけで、ウィンドウ内の一字が同定可能である。カーソルを移動するたびに、新しい子アドレスが自動表示される。

「目次」の例はFig.1の⑧に示した。⑧中の章節番号を発声すると、各章節の先頭行へカーソル移動する機能である。⑧の構造は、テキスト読み込み時に自動的に作成される。

本論文では「索引」の例として、日付と金額を実現した(Fig.1の⑥)。⑥で「日付」を選択すると、日付(〇年〇月〇日、〇年、〇月、〇日、など)の文字列だけを、全テキスト上で

音声 アクセス機構	目的地	ウィンドウ外	
		場所を認知(b)	場所を無知(c)
音声カーソル	ウインドウ内 (a)	○	○
目次機能	△	◎	△
索引機能	△	○	◎

Fig.3 音声アクセス機構の特性

強調表示する。強調表示後は、強調表示した日付だけを飛び飛びにカーソル移動できる。

書籍の索引では書籍の本文中に出現する個々の文字列だけを索引するが、本論文の「索引」ではテキストデータ中に出現する個々の文字列に加えて、文字列のグループ(日付や金額など)を索引できる。本論文では日付と金額しか実現していないが、この他に考えられる「文字列のグループ」として、会社名、商品名、動物名、ソフトウェア名など、その広がりは無限に近い。「索引」は、少ない発声で目的地へ到達し易くするための強力な機構と考える。

Fig.3は、3つの音声アクセス機構を比較している。横軸は、カーソル移動の目的地を三つに分類している。目的地が、ウィンドウ内にある場合(a)、ウィンドウ外でユーザが場所(どの章節にあるのかなど)を知っている場合(b)、ウィンドウ外でユーザが場所を知らない場合(c)である。本論文では、3種類のどの場合でも効率よく指示できるよう、音声アクセス機構を提案した。

4 おわりに

本報告では、音声による指示がキーボードやマウスに比べ難しいことに着目し、少ない発声で指示可能な、新しい音声アクセス機構を提案した。本研究の目次・索引機能から類推できるように、音声指示のメタファは、書籍のそれであることが明らかとなった。現段階は、音声アクセス機構を実現した段階にある。マウス・キー入力を用いた場合における評価実験、実際に音声で制御した場合の評価実験、より適切な発声語彙の検討、などが今後の課題である。

参考文献

- [1] 守屋, 浅貝, 長井, 原, “音声入力テキストエディタにおける指示機構”, 第14回HIシンポジウム論文集, pp.221-230(1998).