

## 文書編集における音声制御の一方式†

中谷 吉久<sup>††</sup> 守屋 慎次<sup>†††</sup>

本論文では、文書の編集作業を対象として音声制御の有効性を調べた。具体的には、(1)編集コマンドを音声で制御する方法を、(2)語彙と構文とに分けて論じた。語彙については、複合名詞（例えば“文字削除”）から成る語彙と名詞句（例えば“文字の削除”）から成る語彙のどちらが発声に向き、より良い認識率が得られるかを実験を通して調べた。構文については、コマンドメニューを逐一選択する方式とショートカットする方式のそれぞれをキーで行う場合と音声で行う場合について実験した。結果として、被験者全員において、ショートカット方式に音声を用いた場合が最も所要時間が短かった。ショートカット方式に音声を用いた場合は、逐一選択方式にキーを用いた場合に比較して平均 21%（被験者によっては 8~30%）の所要時間の削減が見られた。ショートカット方式に音声を用いた場合は、同じショートカット方式にキーを用いた場合に比較して平均 11% の所要時間の削減が見られた。逐一選択方式に音声を用いた場合と同じ逐一選択方式にキーを用いた場合とでは所要時間に統計的な有意差はなかった。また本論文では、どのような編集コマンドを音声コマンドにしたらよいかについて3案を示した。

### 1. ま え が き

本論文の目的は、文書の編集作業において、どのような編集コマンドをどのように発声して入力すれば編集作業がより効率よく行えるかを示すことである。

音声をコマンド入力の手段として用いることには、次の利点が考えられる。

- (1) 指、手、および目の動きを軽減できる。
- (2) 画面上に現れていない“もの”，指令，過去の状態を指示することができる。
- (3) 単独のキーに比較して、より想起の容易な語彙が使用できる。

対話型システムを設計する際にこのような利点を適切に活用する方法が明らかになれば、音声をコマンド入力の手段とすることにより対話型システムの性能の改善が期待できる。それでは、どのような性能の改善が期待できるであろうか。

対話型システムの性能を評価する次元として Card<sup>1)</sup> は 8 項目を挙げているが、その中で数量化が比較的容易な項目を列挙すると、所要時間、誤りの数、習熟の容易さ、思い出しやすさの 4 つがある。

これらを用いて音声入力法を評価した研究として以下のものがある。

Schmandt<sup>2)</sup> は、ウィンドウシステムにおけるウ

ィンドウ操作、例えば「目的とするウィンドウを画面の最前面に表示し、その中にカーソルを移動する」という操作に音声を適用した例を報告している。そこでは 2 か月間にわたって被験者 6 人にそのシステムを使用させ、どのような状況で被験者は音声入力を選択し、どこで問題が生じ、ウィンドウ操作にどのような影響をもたらしたかを調べた。結果として、一部分もしくは全体が隠れているウィンドウに対しては速く操作できたこと、すなわち所要時間が短かったことを報告している。

Morrison<sup>3)</sup> は、キー入力と音声入力を所要時間および誤りの数について比較する実験を行った。その実験は、キー入力に熟練したタイピストとそうでない非タイピストに対して二種類のエディタ（一方はコマンドは簡単だがコマンド数が多く、他方はコマンドは複雑だがコマンド数が少ない）を用いて行われた。結果は、非タイピストがコマンド数の少ないエディタを用いた場合のみ、音声入力の方がキー入力より 1 コマンド当たりの所要時間が短かった。その他の場合は、音声入力の有効性は示されなかった。

Schurick<sup>4)</sup> は文章、数値、およびコマンドを文字単位、単語単位、および句単位で入力する際のキー入力と音声入力をやはり所要時間と誤りの数で比較している。そして、音声入力は単語単位や句単位による文章の入力には有利であるが、ファンクションキーで入力するコマンドや数値の場合は音声入力の利点は見られないと報告している。

本論文では所要時間に焦点ををぼり、文書の編集作業、特に多段のメニューを選択する作業のみを対象としてキー入力と音声入力を比較する。多段のメニュー

† A Method to Control Japanese Word Processor by Voice Input by YOSHIHISA NAKAYA (Computer Section, Technological Administration Division, Industrial Research Institute of Kanagawa Prefecture) and SHINJI MORIYA (Department of Electrical Communications Engineering, Faculty of Engineering, Tokyo Denki University).

†† 神奈川県工業試験所技術管理部電子計算科

††† 東京電機大学工学部電気通信工学科

をキーを用いて選択するためには何度か手を動かす必要があるが、音声の利用によりこれらの動きが軽減され、所要時間の削減につながるものとする。また、音声制御の有効性を調べる際に、どのような語が発声に向くか、および、どのような文脈において音声制御が有効か、に注目する。前者は発声に用いる語彙の問題（以下、発声に用いる語を発声語、その集まりを発声語彙という）、後者は発声語をコマンド体系に組み込む際の構文の問題である。このような視点から論ずることが、発声語をコマンド体系に組み込む際には特に重要であるとする。

まず初めに、本論文で用いる音声認識システムの概要を2章で述べる。3章では、文書の編集作業の構文に着目し、どの部分に音声のコマンド（以下、音声コマンドという）を用いることが有効かについて3例を挙げて説明する。4章では3章で挙げた3例のうちの1例の音声コマンドを実現し、それをを用いて編集作業を行った実験の結果を、語彙と構文の視点から報告する。ここでは、メニュー選択に音声を使うことによって編集に要する時間が8~30%短縮できたことを示す。この結果は、メニュー選択が頻繁で、かつメニュー木が比較的深い場合には、音声制御が有効であることを示唆する。5章では総括的な結論を述べる。

## 2. 音声認識システムの概要

本論文で使用した音声認識システムは、パーソナルコンピュータ(PC 9801)用に開発された市販の音声認識ボード((株)リコー製)<sup>5)</sup>とその制御用ソフトウェア、およびマイクロフォンから成る。この音声認識システムが認識できるのは、あらかじめ登録されている特定話者の発声語彙で、登録可能な数は最大255語である。1語は2秒以内に発声し終える長さでなければならない。なお、この制御用ソフトウェアは、このパーソナルコンピュータ上のワードプロセッサ「一太郎 Ver. 4<sup>6)</sup>」または他のアプリケーションが稼動中にそのキーストローク列をエミュレートするように設計されている。

この音声認識システムの使用方法を以下に一太郎を例にして述べる。

### ●発声語彙の登録法

- ①この制御用ソフトウェアを始動する。
- ②語を3回発声して登録する。その発声語に対応して一太郎が実行すべき動作を、一太郎をキー操作する際のキーストローク列によって定義す

る。

### ●一太郎の実行中における音声コマンドの使用法

- ①一太郎を始動する。
- ②HOME キーを押す。(これにより一太郎のキー入力モードから音声入力モードに切り替わり、この制御用ソフトウェアが始動する。)
- ③発声語を発声する。(これによりその発声語が認識ボードによって認識され、認識結果の第1位の候補が制御用ソフトウェアに渡される。制御用ソフトウェアは第1位の候補に割り当てられたキーストローク列を走査して一太郎を駆動する。駆動し終わると制御が一太郎に戻り、音声入力モードからキー入力モードに切り替わる。すなわち、ある発声語を発声すると、音声認識結果が正しい場合には、その発声語に割り当てられたキーストローク列をキーボードから打鍵した場合と同じ結果を得ることができる。)

### 3. 音声入力向きの編集コマンド、および音声認識システムが備えるべき機能

本章では利用者が文書を編集する際に、どのような編集コマンドを音声で入力すれば、編集作業がより速く行えるかについて論ずる。また、そのような音声入力を有効(または可能)にするために音声認識システムが備えるべき機能について論ずる。対象とした日本語ワードプロセッサは「一太郎 Ver. 4」である。

まず初めに、一太郎における編集コマンドのメニュー木構造に着目し、音声による入力が有効であると思われる編集コマンドを3例挙げ、それぞれの機能およびそれをを用いることの有効性を具体的に説明する。

一太郎では、主な編集作業は多段のメニュー画面上で多数のメニュー項目のうち1つを選択することにより実行される。すなわち、メニュー項目がつくる木構造の上を頻繁に上下左右することになる。それに伴って視線は、カーソルが指す文章部分とメニュー項目とを行き来し、初心者の場合はさらに画面とキーボード間をも行き来する。手もキーボード上を頻繁に動く。したがって、1章で述べた音声の利点を適切に活用するという観点から、例えば次に挙げる3種類の音声コマンドを用意することは有効と考える。

#### (1) メニュー木上のショートカット

ショートカットとは、メニュー木上のある節(node)から、子(や孫)の節を飛び越して子孫の節へ到達す

る操作を指す。ショートカットを発声によって実現することにより、指と手と目の動きが軽減できるので、入力時間の削減が期待できる。さらに、意味のある発声語彙を用いることにより記憶や想起が容易になると予想する。

一例として、文字列の移動を考える。一太郎を用いて文字列を移動するには、通常次の手順に従う。

- ①ESC キーを押す (コマンドメニューが表示される)。
- ②Mキーを押すか、またはメニュー上のカーソルを「移動」の上へ移して改行キーを押す (「移動」コマンドが選択される)。
- ③Cキーを押すか、またはメニュー上のカーソルを「文字」の上へ移して改行キーを押す (「文字」単位で範囲指定することを示す)。
- ④文書上のカーソルを、移動したい文字列の始点へ移して改行キーを押す (始点となる文字が指定される)。
- ⑤文書上のカーソルを、移動したい文字列の終点へ移して改行キーを押す (終点となる文字が指定される)。
- ⑥文書上のカーソルを、文字列の移動先へ移して改行キーを押す (移動先が指定されて移動コマンドが実行される)。

ここで上記①～③の操作は、文字列を移動する際の決まりきった操作である。したがって、ショートカットして差し支えない操作である。例えば「移動」と発声すれば①②③を実行するように登録する。

### (2) メニュー木における兄弟項目間の枝渡り

例えば、罫線コマンドを用いて表を作成する場合、罫線を引いたり消したり、枠を拡大したり縮小したりすることが頻繁に行われる。それには罫線コマンドのメニュー木において、「兄弟」関係にあるメニュー項目 (という枝) 群の上を、「枝渡り (traverse)」することが必要である。ここで木構造における「兄弟」とは同一の節の下につながる枝群のことを指す。ところが一太郎では、この種の枝渡りを1手順で行う方法がない。したがって、この種の枝渡りを音声コマンド (発声は1回) で実現することはかなり有効であると予想する。

### (3) 擬似自然言語を用いた音声コマンド

例えば、(1)で例示した文字列の移動を再び考える。一太郎では(1)で示した手順のほかに、「範囲の先指定」 (例えば(1)の手順①～⑥において、④と⑤

を①の前に行う操作、すなわち移動すべき文字列の範囲を「移動」コマンドを選択する前に指定する操作) と「実行位置の先指定」 (例えば(1)の手順①～⑥において、⑥を①の前に行う操作、すなわち移動先を「移動」コマンドを選択する前に指定する操作) を行うことができる。範囲と実行位置を先に指定し、その後で操作の種類 (例えば「移動」) を指定する方法は、日本語における文節の出現順に沿っている。ところが、一太郎では範囲の先指定と実行位置の先指定にそれぞれ CTRL+] キーと CTRL+^ キーを割り当てている。これを記憶したり思い出すことは利用者には負担であり、思考の流れを乱すものである。そこで、範囲と実行位置を指定した後でコマンドを指定するというこの操作手順を、日本語の音声で“ここから”、“ここまでを”、“ここへ”、“移動”という具合に発声しながら、範囲と実行位置をマウスで指定できるならば、これらの発声語の意味と順序が日本語のそれらと整合して記憶と発声が容易になると同時に、視線は画面上に置いたままで手の動きもごくわずかで済むものと予想する。なお、「擬似自然言語」における発声語彙を上記だけに限る理由はなく、さらに広げることとも可能である。

上記(1)の音声コマンドが有効であることを4章で実験を通して示す。

上記(2)の音声コマンドを実現するためには、発声語に登録するキー列の途中 (キー列の先頭でない位置) に、ESC キーを含める必要がある。しかし、本論文で用いている音声認識システムでは、キー列の途中に ESC キーを含めて登録した場合、その ESC キー以後のキー列が無視されてしまう。したがって上記(2)の音声コマンドは実現できない。

上記(3)の音声コマンドについては、やはり本論文で用いている音声認識システムの次の制約により、その有効性を示すのは難しい。制約とは、キー入力から音声入力へモードを切り換えるたびに HOME キーを押さなければならないことである。この制約により、上記(3)のような音声コマンドを登録して使用したとしても指や目の動きがあまり軽減されないため、この音声認識システムの上ではあまり有効でないと思われる。

以上の理由により、上記(2)、(3)の音声コマンドを用いた実験については本論文ではこれ以上述べないことにする。

以上に述べたことから次のことがわかる。すなわ

ち、音声コマンドが有効そうな（4章で示すように、(1)については有効である）場面がいくつか考えられるにもかかわらず、音声認識する方法（例えばHOMEキー押下が必要）や認識システムを構成する方法（例えばESCキーの問題）自身が、音声入力の可能性を阻んでいる場合がある。したがって、音声入力を利用する立場からの要求を満たすように、認識の方法や認識システムの構成法を研究する必要がある。

#### 4. 音声によるコマンドメニューのショートカット

本章では、3章(1)で述べたコマンドメニューをショートカットする音声コマンドの有効性を、発声する語彙とその構文の視点から実験することによって確認する。

##### 4.1 発声語彙の認識実験

ここでは予備実験として、より良い認識率が得やすい発声語彙とはどのようなものかを実験をとおして調べる。以下にその実験内容を示す。

##### (1) 実験方法

語彙I、IIを用意した（図1）。それぞれは19種類の語から成る。削除、移動、コピーは一太郎の編集コマンドの操作名である。文字、行、段落等は一太郎が用意している操作対象の単位である。語彙Iは操作対象の単位と操作名を連結して複合名詞としたもの

語彙I： 操作対象の単位とコマンド名を連結して複合名詞としたもの（19種類）

文字削除	文字移動	文字コピー
行削除	行移動	行コピー
段落削除	段落移動	段落コピー
ページ削除	ページ移動	ページコピー
ブロック削除	上書ブロック移動	上書ブロックコピー
全行削除	挿入ブロック移動	挿入ブロックコピー
全文書削除		

語彙II： 操作対象の単位とコマンド名を格助詞“の”で区切り名詞句としたもの（19種類）

文字の削除	文字の移動	文字のコピー
行の削除	行の移動	行のコピー
段落の削除	段落の移動	段落のコピー
ページの削除	ページの移動	ページのコピー
ブロックの削除	上書ブロックの移動	上書ブロックのコピー
全行の削除	挿入ブロックの移動	挿入ブロックのコピー
全文書の削除		

図1 複合名詞から成る語彙Iと名詞句から成る語彙II  
Fig. 1 Vocabulary I which is a set of compound words and Vocabulary II which is a set of noun phrases.

で、語彙IIはそれらを格助詞“の”で区切り名詞句としたものである。この種の音声コマンドを実際に用いて編集作業を行った結果については4.2節で述べる。

##### ①語彙の登録

図1に示す語彙I、IIのそれぞれに対して、最左列の上から下へ、次にその右隣りの列の上から下へという順番で登録した。

##### ②認識実験

語彙I、IIのそれぞれに対して各語を6回ずつ、したがって語彙I、IIのそれぞれに対して114回発声して認識実験を行った。IとIIの実験順は不同である。また、114個の語は乱数によって付けられた順に発声した。

被験者は8人で、全員が20歳代前半の男性である。「一太郎 Ver. 4」の使用経験は全員が1年以内で、全く使用経験のない者もいた。被験者8人のうち、2人（A、B）は1989年10月に、残り6人（C～H）は1990年4月にそれぞれ実験を行った。使用したパーソナルコンピュータはPC 9801 RAで、クロック周波数は20 MHzとした。実験は、実験者および被験者を併せて2～3人しかいない比較的静かな個室で行った。

##### (2) 結果

被験者8人について語彙の認識実験を行った結果を図2に示す。図2から次のことが分かった。

①実験を実施した年による認識率への影響は見られなかった。すなわち、被験者A、BとC～Hとの間に統計的な有意差はなかった。したがって、語彙IとIIの認識率については、被験者A、BとC～Hとを1つの母集団として扱うことにした。

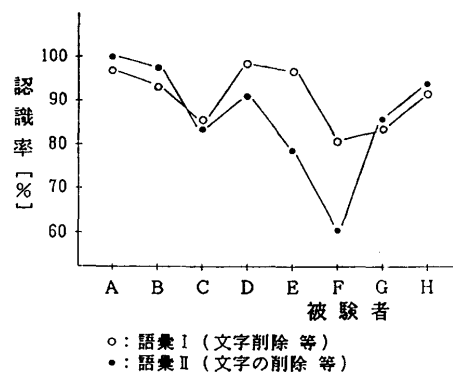


図2 被験者A～Hの語彙I、IIに対する認識率  
Fig. 2 Recognition rates for Vocabulary I and II uttered by each of the eight subjects, A to H. Symbols ○ and ● show recognition rates for Vocabulary I and II respectively.

②被験者8人において、語彙IとIIの認識率に統計的な有意差はなかった。したがってこれらの被験者については、語彙IとIIのどちらを使用してもよいことが分かった。

③被験者ごとに見ると、被験者A, B, CおよびG, Hは語彙IとIIの認識率がほぼ等しくなったが、被験者D, E, Fは語彙Iの認識率の方がかなりよかった。すなわち、語彙IとIIの認識率の差には個人差があった。その理由としては、ほとんどの被験者は語彙Iを好み、語彙IIは発声しにくいと実験後に感想を述べていたことが挙げられる。一般に、発声しにくい語彙は発声が不安定になりやすい。したがって、特に被験者D, E, Fは登録時および認識時に発声が不安定になり認識率が下がったのではないかと考える。

以上の結果から、次に述べる編集作業の実験では、比較的認識率が高く、個人差の少ない語彙Iを使用することとした。

#### 4.2 編集作業実験

本節では、被験者に実際に文書の編集作業を行わせ、音声をどのような作業に対して、すなわちどのような文脈において適用すれば、どのくらいの効果があるかを調べる。

以下に実験方法と結果を示す。

##### (1) 実験方法

被験者は一太郎を用いて日本語文書(論文)の編集作業を行う。編集作業の内容を被験者に指示する用紙はB4判で20頁から成る。その一部を図3に示す。図示したように、文章データはすでに入力されている。被験者が行う作業は、図3の用紙を見ながら文章データの一部を削除または移動またはコピーすることである。途中で文字を入力することはない。編集作業は合計27個である。その27作業の内訳は、3種類の単位(文字, 行, 段落)のそれぞれに対して3種類の操作(削除, 移動, コピー)を各々3回ずつ行うことである。ただし、27作業は文書の中に順不同であらわれる。

被験者は前述の実験を行った8人と同一である。そのうち2人(A, B)は1989年10月に、残り6人(C~H)は1990年4月にそれぞれ実験を行った。発

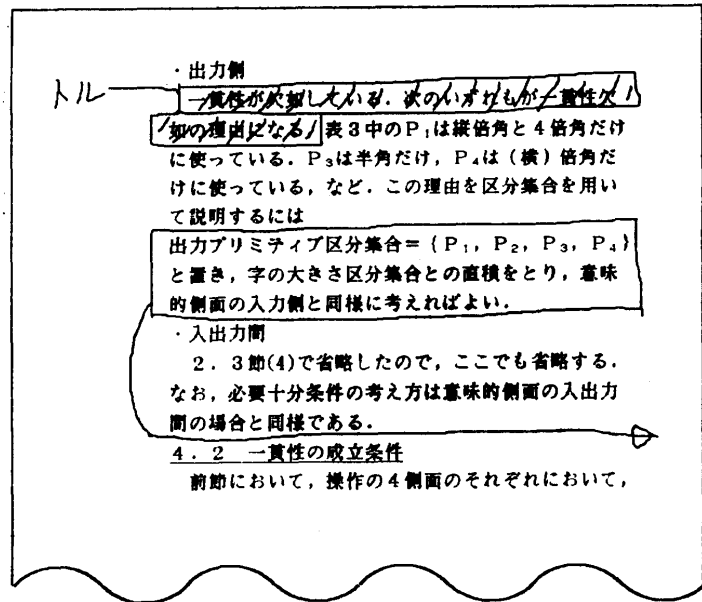


図3 編集作業を指示する用紙の一部

Fig. 3 A part of sheets that instruct each of the eight subjects how to edit a document.

声語彙は被験者全員が語彙I(図1)を用いた。使用したパーソナルコンピュータおよびクロック周波数は前述の実験と同一である。

被験者はコマンドメニューの選択法として次の2つを実施した。

①逐一選択方式(メニュー木上の節を逐一、選択していく、すなわち画面上に見えているメニュー項目を逐一、選択していく方式)。

②ショートカット方式(メニュー木上のいくつかの節をショートカットする、すなわち画面上に見えていない何段階か先の状態を指示する方式)。

この2方式の各々に対して使用できる入力機器は次の2通りとした。

①キーボードとマウスのみ(ただし、逐一選択方式におけるメニューの選択とショートカット方式におけるショートカットの指示はキーで行う)。

②キーボードとマウスと音声(ただし、逐一選択方式におけるメニューの選択とショートカット方式におけるショートカットの指示は音声で行う。しかし、誤認識が3回連続した場合はキーを使うものとする)。

したがって、各々の被験者は27個の編集作業を、2種類のメニュー選択法とその各々に対する2通りの入力機器の合計4種類の入力方法で行った。

実験の順序は、逐一選択方式とショートカット方式

のうち、前者を最初に行った者が6人（そのうちキーを最初に用いた者が3人、音声で最初に用いた者が3人）、後者を最初に行った者が2人（そのうちキーを最初に用いた者が1人、音声で最初に用いた者が1人）である。

編集作業の所要時間は、実験に同席した測定者がストップウォッチで測定した。ここで「編集作業の所要時間」とは、測定者が被験者に「はじめ」の合図を与えてから、被験者が図3の用紙上に作業の指示がもうないことを確認するまでの時間をいう。

実験は、実験者（測定者を兼ねる場合もある）、測定者、および被験者を併せて2~3人しかいない比較的静かな個室で行った。

なお、ショートカット方式に音声を用いる場合の発声語彙とそれへのキー列の登録内容を図4に示す。図4の左辺に示された語彙を発声すると、右辺に示されたキー列がエミュレートされる。同様に逐一選択方式に音声を用いる場合の発声語彙およびキー列の登録内容を図5に、またショートカット方式にキーを用いる

発声する語彙		登録するキー列
文字削除	←	ESC D C
行削除	←	ESC D L
段落削除	←	ESC D S
文字移動	←	ESC M C
行移動	←	ESC M L
段落移動	←	ESC M S
文字コピー	←	ESC C C
行コピー	←	ESC C L
段落コピー	←	ESC C S

図4 ショートカット方式に音声を用いる場合の発声する語彙とその語彙が発声されたときにエミュレートされるキー列

Fig. 4 Each line shows a vocal word and a key string which is emulated when a subject utters the vocal word. These vocal words are used to short-cut menu selections by speech recognition.

発声する語彙		登録するキー
エスケープ	←	ESC
削除	←	D
移動	←	M
コピー	←	C
文字	←	C
行	←	L
段落	←	S

図5 逐一選択方式に音声を用いる場合の発声する語彙とその語彙が発声されたときにエミュレートされるキー

Fig. 5 Each line shows a vocal word and a key which is emulated when a subject utters the vocal word. These vocal words are used to select a menu item by speech recognition.

入力する語彙		登録するキー列
F1	←	ESC D C
F2	←	ESC D L
F3	←	ESC D S
F4	←	ESC M C
F5	←	ESC M L
CTRL + F6	←	ESC M S
CTRL + F7	←	ESC C C
CTRL + F8	←	ESC C L
CTRL + F9	←	ESC C S

図6 ショートカット方式にキーを用いる場合の入力語彙とその語彙が入力されたときにエミュレートされるキー列

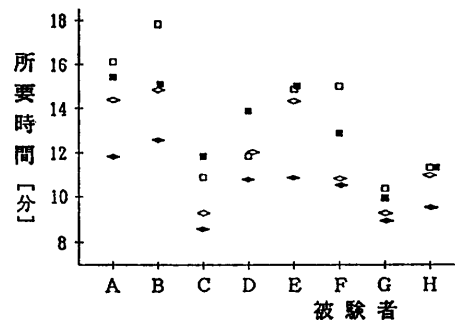
Fig. 6 Each line shows a function key and a key string which is emulated when a subject inputs the function key. These function keys are used to short-cut menu selections.

場合の語彙とそれへのキー列の登録内容を図6にそれぞれ示す。逐一選択方式にキーを用いる場合、登録するキー列はない。

(2) 結果

被験者8人（A~H）について、編集作業実験を行った結果を図7に示す。図7から次のことが分かった。

①被験者8人全員において、ショートカット方式に音声を用いた場合が最も所要時間が短かった。これは、各被験者が行う4種類の入力方法の順序に依存



□: 逐一選択方式にキーを使用  
 ■: 逐一選択方式に音声を使用  
 ◇: ショートカット方式にキーを使用  
 ◆: ショートカット方式に音声を使用

図7 被験者A~Hの編集作業の所要時間

Fig. 7 Time required by each of the subjects, A to H, when editing a document. The four symbols: □ is for the step-by-step selection technique with a keyboard and mouse; ■ is for the step-by-step selection technique with a keyboard, a mouse, and the speech recognition device; ◇ is for the short-cutting menu selection technique with a keyboard and mouse; ◆ is for the short-cutting menu selection technique with a keyboard, a mouse, and the speech recognition device.

していない。

②実験を実施した年による所要時間への影響が見られた。すなわち、被験者A、BとC～Hとの間に統計的な有意差があった（危険率5%）。これは、実験内容への慣れおよび一太郎に対する熟練度等が影響したと考えられる。

以下に述べる議論③④⑤において、被験者A、BとC～Hを1つの母集団とした場合とそれぞれを別の母集団とした場合とで統計的な処理を施したところ、いずれの場合も有意差の有無については同じ結果が得られた（ただし、統計的な有意差の危険率が異なる場合はある）。したがって、以下では被験者A、BとC～Hとを1つの母集団とした場合の結果のみを述べることにする。

③逐一選択方式にキーを用いた場合と同じ逐一選択方式に音声を用いた場合とでは、所要時間に統計的な有意差はなかった。その理由を次のように考える。画面上に見えているメニューの項目をキーで指示する場合は1手順（英字1字またはカナ1字の入力）であり、音声で指示する場合は2手順（HOMEキーを押した後に発声語を発声する）である。すなわち、音声を使用しても入力の手順数が減らず、目や手の動きが軽減されない。このように、1章の冒頭で挙げた音声の利点が活用されていないため、音声とキーとで差がでなかった。

④ショートカット方式に音声を用いた場合は、逐一選択方式にキーを用いた場合と比較して平均21.2%（被験者によって約8～30%の幅がある）の所要時間の削減が見られた。また、これらの間には統計的な有意差のあることが分かった（危険率1%）。その理由を次のように考える。メニューのショートカットを音声で指示する場合は2手順（上記③と同様）を必要とするが、一方そのショートカットと同等な作業をメニューを逐一選択して行う場合には3手順が必要である（例えば、文字列の移動にはESC、M、Cキーの打鍵）。すなわち、画面上に見えていないもの（状態）が音声を用いることにより少ない手順数で指示でき、目や手の動きが軽減できたため、所要時間が削減できた。

⑤ショートカット方式に音声を用いた場合は同じショートカット方式にキーを用いた場合より11.4%の所要時間の削減が見られた。また、これらの間には統計的な有意差のあることが分かった（危険率1%）。その理由を次のように考える。メニューの

表1 被験者A～Hの実験後における感想

Table 1 A written description of the subjects' impressions after editing a document with a keyboard or the speech recognition device.

	[キー, マウス]	[キー, マウス, 音声]
長所	・確実に実行される。	・キー入力よりも速い。
短所	・コマンドを画面上とキーボード上で探すのが面倒だ。 ・コマンドを登録したキーを憶える必要がある（ショートカット方式）。	・正しく認識したかどうかの確認がわずらわしい。 ・誤認識した時にやり直すのが面倒だ。 ・使用する時にどんな発声語であったかを考え出す必要がある（ショートカット方式）。

ショートカットをキーで指示する場合は1手順（ファンクションキーの押下）で済み、音声で指示する場合より手順数が少ない。しかし、音声の場合は作業内容から発声語が容易に連想できるのに対して、キーの場合はファンクションキーしか使用できない。すなわち、音声の利用により語彙の想起に要する時間が削減できた。

被験者に編集作業の実験後、感想を述べてもらった。それらを表1に示す。その中で特に重要な音声入力の手順を以下に列挙する。

- (1) 正しく認識したかどうかの確認がわずらわしい。
- (2) 誤認識した時にやり直すのが面倒だ。

音声入力により確かに編集作業の所要時間は短縮されたが、被験者が指摘している上記(1)、(2)の音声入力の手順は重要である。これらについての検討は今後の課題としたい。

## 5. 結 論

本論文では、次のことを示した。

- (1) 文書の編集作業における音声制御の有効性を語彙と構文の視点から示した。語彙については、「複合名詞と名詞句のどちらがより良い認識率が得られるか」という予備実験の結果を示したにすぎない。しかし、利用者に整合した発声語彙を同定すること、仕事に整合した語彙を同定すること、それらの語彙を正しく認識するアルゴリズムを開発することの、それぞれの重要性を示唆したと考える。構文については所要時間に焦点を当て、コマンドメニューのショートカットを音声で制御することの有効性を実験によ

って示した。

- (2) 一太郎を用いて文書を編集する作業に実際に音声を用いて実験を行った。その結果、次のことが分かった。

- ①被験者 8 人に文書の編集作業を 4 種類のコマンド入力法（コマンドメニューを逐一選択する方式とショートカットする方式のそれぞれをキーまたは音声で行う）で行わせた。結果として、被験者全員において、ショートカット方式に音声を用いた場合が最も所要時間が短かった。
  - ②ショートカット方式に音声を用いた場合は、逐一選択方式にキーを用いた場合に比較して平均 21.2%（被験者によっては 8~30%）の所要時間の削減が見られ、また統計的にも有意差のあることが分かった。
  - ③ショートカット方式に音声を用いた場合は、同じショートカット方式にキーを用いた場合に比較して平均 11.4% の所要時間の削減が見られ、また統計的にも有意差のあることが分かった。
  - ④逐一選択方式に音声を用いた場合と同じ逐一選択方式にキーを用いた場合とでは所要時間に統計的な有意差はなかった。
- (3) 音声認識の方法や認識システムの構成法そのものが音声入力の可能性を阻んでいる場合があることを指摘し、認識システムへの要件を例示した。それらを以下に述べる。

- ①発声語に登録するキー列がいかなる場合でも制約を受けないようにする。
- ②キー入力から音声入力へモードを切り換えるための操作を不要にする。

これに関連して次のような提案を行った。すなわち、どのような編集コマンドを音声コマンドにしたらよいかについて 3 案を示した（3 章）。

上記(2)②が得られた理由は、画面上に見えていない指令が音声により少ない手順数で指示でき、目や手の動きが軽減できたためと考える。上記(2)③が得られた理由は、音声はキーに比較してより想起の容易な語彙が使用でき、したがってその語彙の想起に要する時間が削減できたためと考える。

これらは 1 章の冒頭で挙げた音声の利点(1)~(3)が活用できた結果であると言える。言い換えると、上記(2)②③の結果は音声に前述の利点(1)~(3)があることの裏づけのひとつとなっている。Schmandt<sup>2)</sup>

の実験結果も音声の利点(1), (2)を裏づけている。

しかし、上記の結論(2)では登録に要する時間は考慮していない。また、コマンドメニューの木が比較的浅い場合については述べていない。さらに、文書の編集作業全体を考えたとき、一部の編集コマンドに対してのみ音声コマンドを用意すると、利用者が音声コマンドを使用する際に混乱を招く可能性がある。これらに関するより一般的な議論は今後の検討課題としたい。

本論文の結果は音声による制御方式の一部を述べたにすぎないが、その全容を明らかにしていく上で、このような実験を積み重ねることが役立つものと考えられる。

**謝辞** 本研究を進めるにあたり、被験者として快く実験に協力してくださった東京電機大学平成元年度および 2 年度の卒研生諸氏に感謝いたします。

### 参 考 文 献

- 1) Card, S. K., Moran, T. P. and Newell, A.: The Keystroke-Level Model for User Performance Time with Interactive Systems, *Comm. ACM*, Vol. 23, pp. 396-410 (1980).
- 2) Schmandt, C., Ackerman, M. S. and Hindus, D.: Augmenting a Window System with Speech Input, *IEEE Comput.*, Vol. 23, No. 8, pp. 50-56 (1990).
- 3) Morrison, D. L., Green, T. R. G., Shaw, A. C. and Payne, S. J.: Speech-controlled Text-editing: Effects of Input Modality and of Command Structure, *Int. J. Man-Mach. Stud.*, Vol. 21, pp. 49-63 (1984).
- 4) Schurick, J. M.: Efficiency of Limited Vocabulary Speech Recognition for Data Entry Tasks, *Proceedings of the Human Factors Society—30th Annual Meeting*—, pp. 931-935 (1986).
- 5) (株)リコー: NEC PC 9801 用音声認識ボード取扱説明書 (Jan. 1990).
- 6) (株)ジャストシステム: 一太郎 Ver. 4 NEC PC-9800 シリーズ 解説編 (May 1989).

(平成 2 年 11 月 28 日受付)

(平成 3 年 12 月 9 日採録)



**中谷 吉久 (正会員)**

昭和34年生。昭和57年東京電機大学工学部電気通信工学科卒業。昭和63年同大学院博士課程満期退学。現在、神奈川県工業試験所技術管理部電子計算科勤務。対話型システムにおけるユーザインタフェースの分析・検査・評価法および音声入力装置を用いたユーザインタフェースの研究に従事。電子情報通信学会、計測自動制御学会ヒューマン・インタフェース部会各会員。

**守屋 慎次 (正会員)**

昭和48年東京電機大学大学院博士課程修了。工学博士。現在、東京電機大学工学部電気通信工学科教授。昭和56年ニューヨーク州立大学、昭和57年イリノイ大学の各計算機科学科客員準教授。ユーザインタフェース、パターン認識、ペン入力システム、グループウェア、リ音声入力技法、インタラクションの分析・評価・モデリング法の研究に従事。Interacting with Computers誌の Special Editorial Board。電子情報通信学会、計測自動制御学会、人間工学会、テレビジョン学会、ACM, IEEE 各会員。