

# 汎用形音声認識装置“HR-100 II”

## General-use Speech Recognition Equipment “HR-100 II”

音声は、人間の最も自然な情報手段である。この音声情報を情報処理装置への入力手段として利用することは、他の入力装置に比べ優れた点がある。しかし、実際に音声入力を応用する場合、音声応用という性質上、音声入力に関する複雑な手順処理プログラムと応用業務の実行処理プログラムを開発する必要がある。

本装置は、音声認識装置をシステム製品化し、ユーザー負担低減の実現を可能にする支援ソフトウェアを具備している。

システムの操作は、音声入力作業の性質を配慮し、実時間会話形による処理を基本とした。また、ソフトウェア作成時の記述作業を容易化するための、音声入力問題向き言語などの支援ソフトウェア群を開発した。

木村俊宏\* *Toshihiro Kimura*  
 三崎良典\* *Yoshinori Misaki*  
 小松昭男\*\* *Akio Komatsu*  
 姫路正明\*\*\* *Masaaki Himeji*  
 秋本文之\*\*\*\* *Fumiyuki Akimoto*

### 1 緒 言

音声は、人間の最も自然な情報通信手段である。情報処理装置への、データ入力手段として音声を使うことにより、下記のように他の入力手段と比較し、有利な点が多々ある。

#### 1.1 音声入力の利点

- (1) 音声は人間にとって自然な動作であり、特別な訓練なしに使用できる。
- (2) 目や手を他の作業に拘束されているときでも、データを入力するために、その作業を中断することなく同時処理ができる。
- (3) 手袋をして作業する現場、手を汚さなくてはならない現場、手を汚してはいけないような場合でも、データを入力することができる。
- (4) 音声で入力した情報は、直接データ処理装置に入力でき、カード、紙テープへのパンチをしたり、タイプライタをたたくななどの中間処理が不要となる。
- (5) 高所、機械の中などのような狭所、暗室などのような特殊な作業環境の中でも、情報の入力が可能である。
- (6) マイクロホンの届く範囲内であれば、入力装置の設置場所にかかわらず、作業者は移動して入力できる。

以上述べたような利点に着目して、音声認識装置の導入による合理化、省力化をねらって、各方面で装置の導入が計画されている。

しかし、一方、ユーザーが実際に導入する場合に、下記のような問題があり、この対策のため、音声認識装置を制御する技術とソフトウェアの開発負担が伴って、音声入力を普及するひとつの障害にもなっている。

#### 1.2 音声入力装置導入時の問題

まず第1の問題点は、現在の音声認識装置では、ある程度の誤認識は避けられない。この対策のため複雑な手順処理を行なうソフトウェアの開発を必要とすること。

次に第2の問題点は、音声認識装置を利用する場合に、その利点を更に引き伸ばす手段として、作業者の音声入力と、作業への指示、確認とを会話形式で実時間応答処理のシステムにすることを必要とすること。

そして、第3の問題点は、音声入力装置導入の目的は、音声認識そのものではなく、最終的にはデータの編集、編集後の

データ印字出力やホストコンピュータへのデータ渡しである。そのためのソフトウェア開発を必要とすることである。

以上述べた問題解決のため、HR-100 II 汎用形音声認識装置(以下、HR-100 II と略す。)は次に述べるような点を目的に開発を行なった。

#### 1.3 HR-100 II 開発のねらい

まず、音声入力装置をシステム製品化し、ユーザーが本装置を導入すれば、あまり手をかけずにそのまま運用できるようにした。音声入力処理の簡易化、効率化を実現して、その運用方法は、極力操作者が扱いやすいものとした。そして、具体的な応用先としては一般産業向けとし、中でも特に生産技術分野向けを考慮して開発を行なった。

### 2 方 式

音声認識方式と日立製作所の製品名称は、発声する話者、発声条件により表1に示すとおりである。

本報告の認識装置は、話者が登録した本人の認識を対象とする特定話者用、前後に区切って発声された単語認識を対象とする離散単語認識の領域に分類される。

HR-100 II は、先に日立製作所で開発したHR-100 音声認識装置(認識だけを目的としたスタンドアロンタイプ)を使用し、更に、システム操作に必要な入出力機器と、これらを制御する制御部、及びシステムソフトウェアを備えたものである。

### 3 特 長

HR-100 II の特長を下記に、またシステム構成を図1に示す。

表1 音声認識方式と製品名称 認識装置の分類は、発声する話者により二分され、更に、離散発声か連続発声かによって分類される。

話者 発声条件	話 者			
	特 定 話 者		不 特 定 話 者	
離 散 発 声	品名	HR-100/HR-100 II	品名	HR-150
連 続 発 声	品名	HR-200	品名	HR-250

注：品名は、日立製作所の製品名称を示す。

\* 日立製作所戸塚工場 \*\* 日立製作所中央研究所 \*\*\* 日立製作所生産技術研究所 \*\*\*\* 日立製作所システム研究所

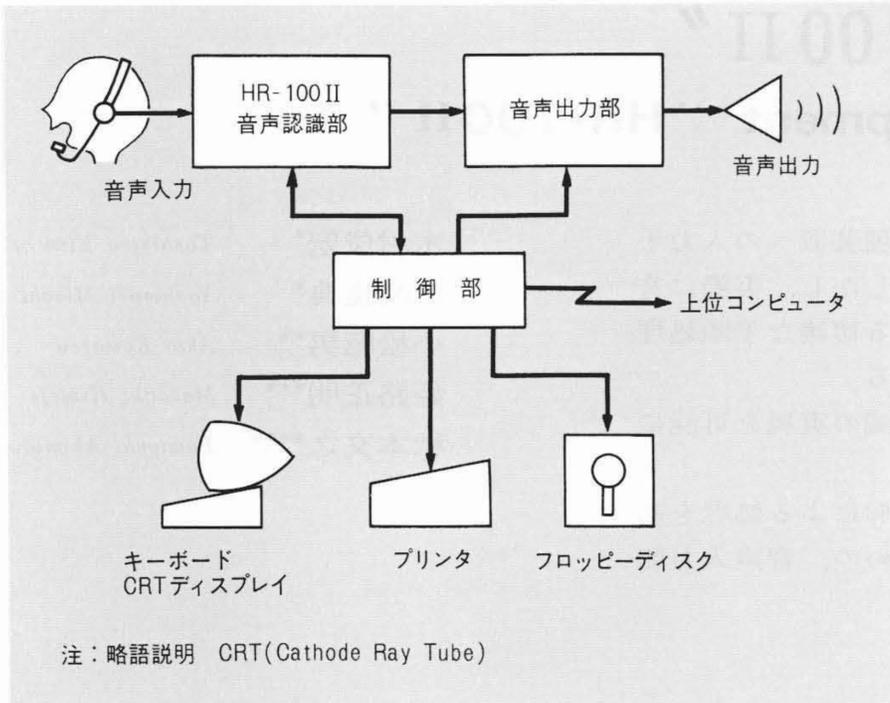


図1 HR-100II 音声認識装置システム構成 キーボード、CRTディスプレイ及び音声出力部は、会話形処理での操作に、フロッピーディスクは、プログラム、データ及び音声標準パターンの格納用に、また、プリンタは、各種データの記録用に使用する。

- (1) 入力される単語の範囲を、あらかじめ予測して認識させる構文的な処理機能を持ち、認識率の改善を実現した。また、音声入力問題向き言語を備え、認識語群の定義や入力構文の記述を簡易化した。更にまた、実機によるプログラム開発を可能とした。これにより、ユーザープログラム開発期間を大幅に短縮できるようにした。
- (2) 操作は、CRT (Cathode Ray Tube) ディスプレイと音声応答による実時間実行の会話形式を基本とし、ユーザーの作業性向上に配慮した。
- (3) 音声登録、認識テストに対し、CRT画面会話形式による登録、格納、テストなどの操作を簡単なキーボード操作により可能とした。
- (4) 各種の応用ユーティリティを備え、会話形レポート編集、会話形ファイル作成、フロッピーディスクフォーマットをホストコンピュータに合わせた変換など、ユーザーの支援業務を可能とした。

次に、以上述べたようなHR-100 II の機能仕様、特長などを、ハードウェア、ソフトウェアについて更に詳しく述べる。

#### 4 ハードウェア構成

##### 4.1 認識部機能ブロック図と動作

HR-100 II 認識部の機能ブロック図を、図2に示す。同図の

左側から入力された音声は、分析部の入力レベルに合った大きさに増幅される。分析部では、入力された音声情報をデジタルフィルタバンクにより、スペクトル分析を行なう。分析されたデータは、このままでは情報量が非常に大きいので、圧縮部で特徴が失わないようにして、情報の圧縮が行なわれる。音声区間検出部では、分析部の途中の情報を利用して、音声の始まりと終わりを区別して次の処理に利用される。

最初の登録時、圧縮されたデータは、標準パターンメモリに格納される。また、通常の認識動作時には、このデータは認識部へ送られ、標準パターンメモリに格納されている標準パターンデータと次々に比較され、その類似度が計算される。判定出力部では、計算された類似度を判定し、最大類似度の単語(番号)を制御部へ送る。

##### 4.2 HR-100 II の仕様

HR-100 II の仕様を表2に示す。音声出力部には、PARCOR合成を使った音声合成パッケージを使用している。音声出力語数及び長さは、合計の語長が100秒以内であれば、ユーザーの希望に応じて自由に選択できる。フロッピーディスクは、標準形(メモリ容量256kバイト)が2台実装されており、1台は装置のシステムプログラムや音声標準パターンの格納用に使用され、他の1台はユーザー業務用ファイルとして開放される。最初に本装置を使用するとき音声に登録し、フロッピーディスクに格納しておけば、そのディスクを差し替えることにより、使用すること登録しなくて済む。キーボードは、作業者が本装置を操作するとき、プログラムの作成

表2 HR-100 II 音声認識装置の主な仕様 本表に示した入出力装置以外にも、拡張用ボード挿入により、ユーザー希望の入出力装置の接続が可能である。

項番	項目	仕様	
1	話者条件	特定話者	
2	認識単位	離散単語	
3	認識語数	32~128語	
4	単語長	0.2~2.0秒	
5	語間長	0.3秒以上	
6	登録時発声回数	1回	
7	認識速度	0.3秒以下(語間長を除く)	
8	入出力装置	音声出力部	PARCOR合成 最大約100秒語
9		フロッピーディスク	標準形 容量 256kバイト 2台
10		キーボード	JIS配列準拠 87キー
11		CRTディスプレイ	キャラクタディスプレイ(80字×25行)
12		プリンタ	速度 80字/秒 1桁 80文字
13	マイクロホン	ダイナミック接話形単一指向性	

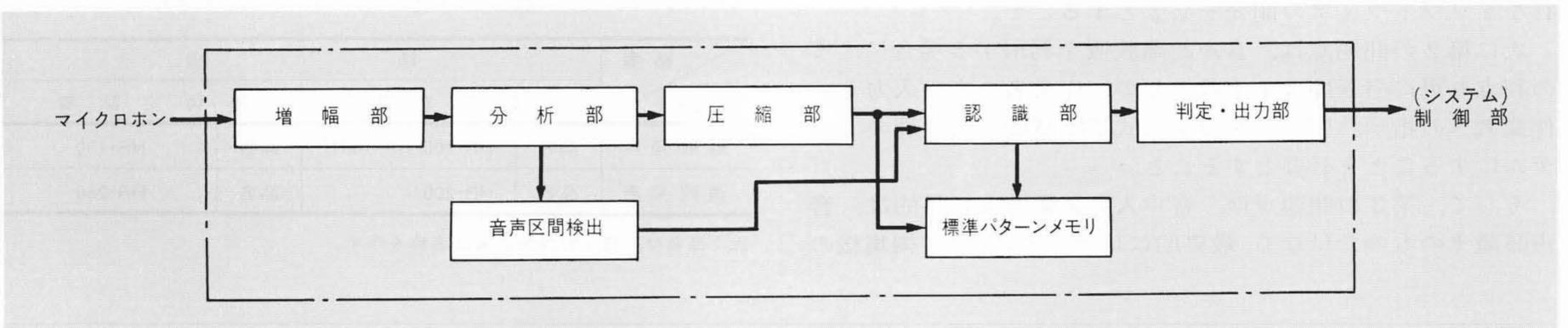


図2 HR-100 II 認識部機能ブロック図 入力された音声は、左から順に右へと処理され、標準パターンとの照合を行なって判定部で認識結果を出力する。

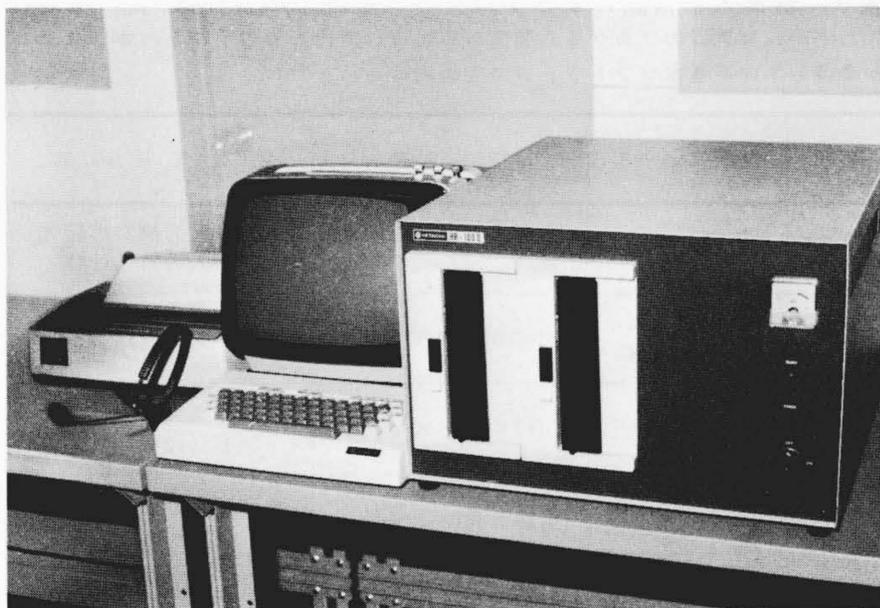


図3 HR-100 II 音声認識装置の外観 右側のフロッピーディスクが実装されているのが本体で、その左側がキーボードとディスプレイであり、左側のものはシリアルプリンタとマイクロホンである。

又は表示メッセージの入力文作成のときなどに使用する。CRTディスプレイは、作業者と本装置の会話形式による操作の中心になるもので、認識結果の表示、各種メッセージの表示に使用される。プリンタは、音声で入力されたデータの編集結果、ユーザープログラム作成時、プログラムリストの出力などに使用される。本装置は、入出力装置拡張用インタフェースを備えているので、拡張用ボードを挿入するだけでユーザーの希望に応じて、他の表示機器、印刷装置、データ入出力

装置への接続が可能である。HR-100 IIの外観を図3に示す。同図右部の筐体が本体で、2台のフロッピーディスクが搭載されている。本体の左側にあるのが、CRTディスプレイとキーボードである。左部にあるものがシリアルプリンタで、その前に置いてあるものがマイクロホンである。

## 5 ソフトウェア構成

### 5.1 概要

従来のような音声認識装置(HR-100)単体では、音声認識とその簡単なサポート機能に限られる。このため、ユーザーが音声入力を実際の省力化システムなどに応用しようとする場合、音声入力に関する複雑な手順を処理する制御プログラムや、応用業務の実行処理プログラムを開発しなければならない。HR-100 IIは、支援ソフトウェアを具備し、ユーザーのソフトウェア開発負担の低減をねらったシステム化製品である。

HR-100 IIソフトウェアの特長は、次に述べるとおりである。

- (1) 本システムの操作は、音声応用という性質上、人間の自然発声速度に追従した会話形実時間処理を基本として行なえるようにした。
- (2) 応用業務用ユーザープログラム作成時、音声入力手順と処理手続などの記述を容易に行なえるよう配慮した。
- (3) 応用ユーティリティは、これを利用すれば、ユーザーは簡単なパラメータ入力とすることにより、本装置をそのままユーザーシステムに応用できるようにした。

図4に、上述の支援ソフトウェアの利用範囲によるユーザーのプログラム開発負担の程度を示す。

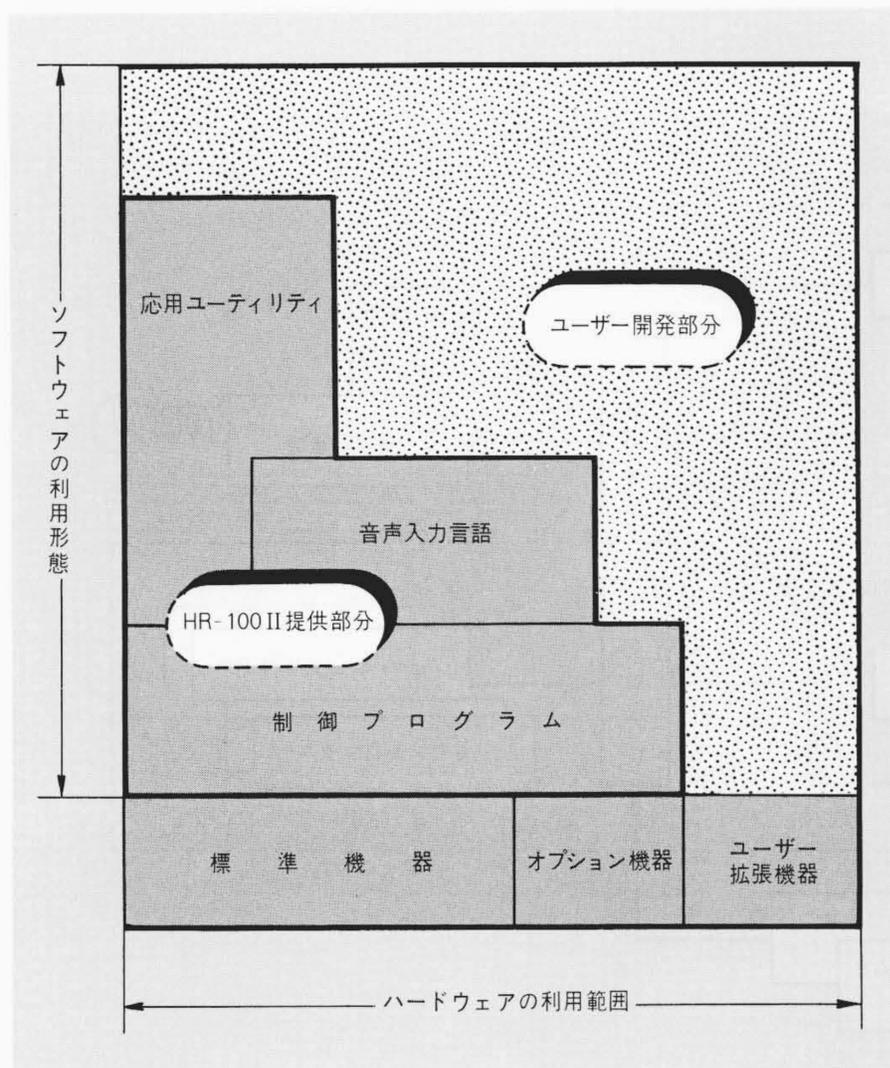


図4 支援ソフトウェアとユーザー開発ソフトウェア HR-100 IIを利用する場合、使用する支援ソフトウェアとユーザーが開発しなければならないソフトウェアの関係を示す(制御プログラム、音声入力言語、応用ユーティリティに対して、右上がユーザー開発部分である)。

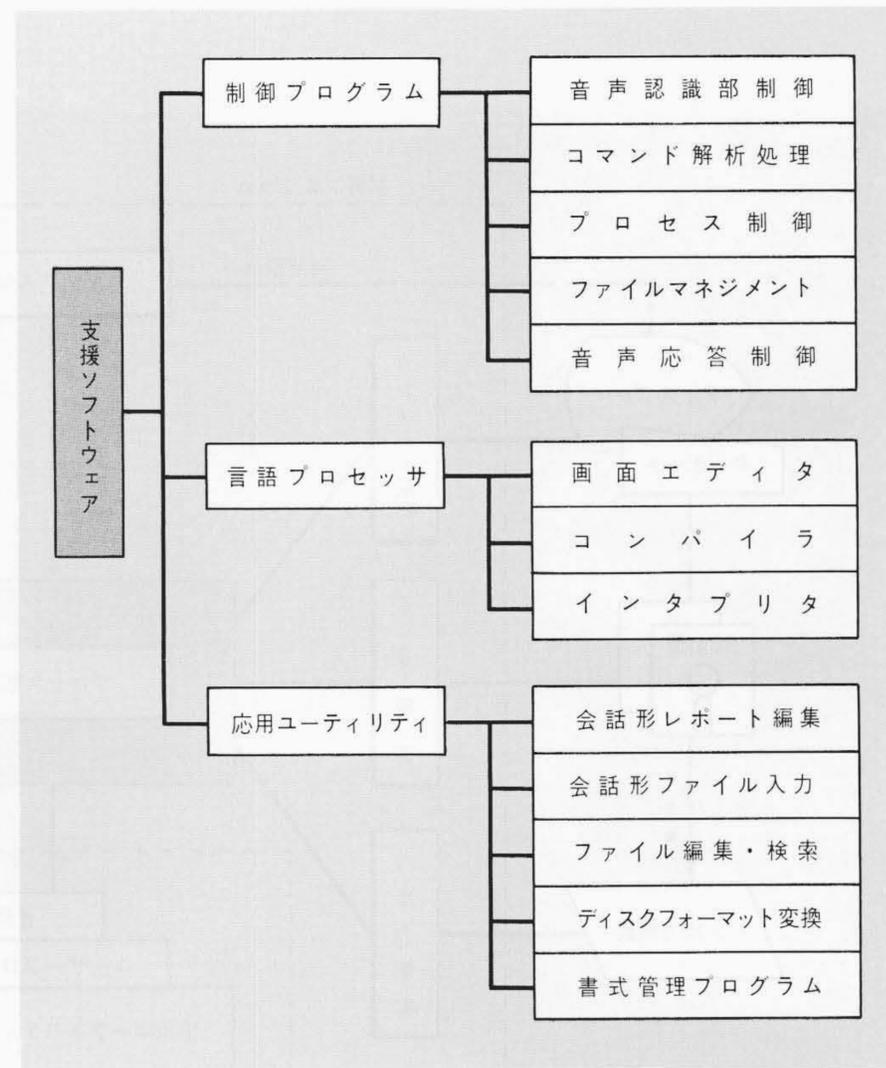


図5 支援ソフトウェアの内容 支援ソフトウェアの構成は、三つのプログラム群から成る。入出力機器及びプログラムの管理を行なう制御プログラム、音声入力用問題向き言語及び応用ユーティリティがあり、後者の二つは、HR-100 IIで特に配慮したものである。



構文の記述をも簡易化した点にある。この言語を利用することにより、構文的な予測認識を効率よく行なえるとともに、プログラム開発期間を大幅に短縮できる。

この音声入力言語を用いて、音声応用プログラムの開発から実際に応用業務を実行するまでの手順、支援内容を表3に示す。同表では本システムでサポートする基本ソフトウェアの各開発フェーズでの支援機能を併記した。各処理過程での操作方法は、後述するので省略する。音声入力言語コンパイラによる中間語形式の変換出力を、インタプリタによって解釈実行する方式を採用している。音声入力特有の誤発声、誤認識のようなあらかじめ予想できない処理変更などの処理をインタプリタにもたせ、ユーザーの音声入力言語の記述を大幅に単純化している。

5.4 応用ユーティリティ

本ユーティリティは、ユーザーが本装置を導入する場合最もユーザー負担の少ない利用形態を提供するプログラム群である。その特長は、以下に述べるとおりである。

(1) 会話形データ入力形式

CRTディスプレイ画面と音声出力とを併用し、視聴覚を組み合わせたデータ入力が可能である。従来のCRTキーボード入力処理形態に比べ、操作性の高い会話形を実現した。

(2) 豊富な支援機能の具備

7種類のデータ処理が用意されており、それらの関係は図7に示すとおりである。レポート編集、ファイル入力作業は、会話形式によるデータ入力方式が利用できる。

(3) ユーザー処理定義の簡易化

ユーザーは、会話形データ入力の会話手順とデータ編集の書式を簡単なパラメータで記述すればよい。パラメータの定義項目の内訳を表4に示す。

表4 応用ユーティリティのユーザー定義項目 本応用ユーティリティプログラム群は、表に示される項目に対応した内容を定義することにより、汎用的にユーザーの業務に適した手順で利用することができる。

会話手順定義		書式定義	
項目	内容	項目	内容
ガイド表示	ガイドメッセージのCRT画面への出力内容	メッセージ表示	見出し文、注釈文の印字法
ガイド音声出力	音声入力の促し、結果判定などの音声出力法	ファイルデータ入力	編集するデータの補助記憶装置から読出し法
ガイド繰返し制御	同一会話手順の繰返し範囲	データ編集出力	データの編集、出力の方法
会話入力	音声入力受けと入力処理法	出力計数	出力ページ数、出力項目数の計数及び表示法
データ格納	入力結果を補助記憶装置への格納法	出力繰返し制御	書式の繰返し利用範囲

(4) 上位計算機との接続の容易化

上位計算機とオフラインで接続する場合でも、フロッピーディスクを介して容易に結合することができる。

5.5 操作機能

HR-100 II システムのコンソール操作では、次の二つの処理モードから構成される。

(1) 主操作モード

システム起動によって本モードに入り、コマンド入力待ちとなる。プログラムの開発、実行はすべてこのモード下である。この処理操作で使用できるコマンドの項目と内容を、表5に示す。

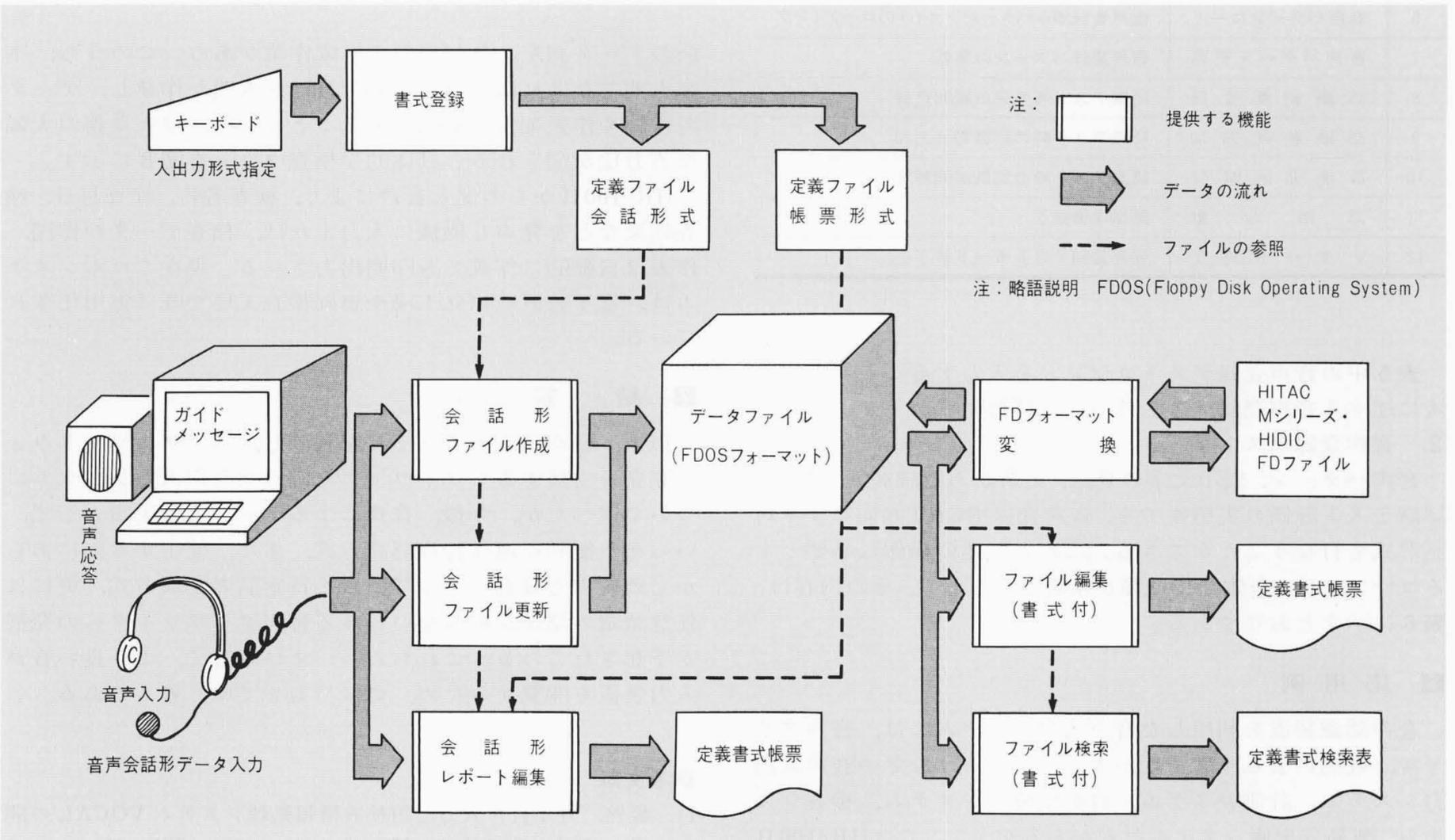


図7 応用ユーティリティの支援業務 HR-100IIの標準ソフトウェアで、図4の最もユーザー負担の少ない利用形態を提供するためのプログラム群である。支援する業務の範囲と各プログラム処理の関連を示した。

表5 主操作コマンド一覧 本表はシステム操作用のコマンドを示し、CRTキーボードから、対話的に入力することによりHR-100IIシステムの操作、運用ができる。

No.	コマンド名称	内 容
1	プロセス*起動	プログラムの実行要求
2	プロセス停止	プログラムの停止要求
3	メモリダンプ	主記憶部の内容を画面表示
4	メモリ内容変更	主記憶部の内容を変更する。
5	CRTプリント	CRT画面内容をプリントする。
6	ディレクトリ表示	ファイルディレクトリの画面表示
7	ファイル入力	プログラム、データファイルのメモリロード
8	ファイル生成	プログラム、データファイルのフォーマット生成
9	音声登録テスト	音声認識テストモードへの切替
10	エディタ起動	画面エディタのメモリロード実行
11	コンパイラ起動	音声入力言語コンパイラのロード実行
12	インタプリタ起動	音声入力インタプリタのロード実行

注：\* 一つの独立したプログラム単位であり、各プロセスに対応したプロセス番号(最大4個)が割当てできる。

表6 音声登録コマンド一覧表 本表は、表5の音声登録テスト下でのコンソール入力コマンド群を示しており、音声登録、更新、学習、認識テストなどを行なうときに用いる。

No.	コマンド名称	実 行 内 容
1	初期設定	音声認識部の初期処理実施
2	文字登録	CRT画面上に語番、発声文字の表示と更新
3	文字ストア	音声登録画面のFDファイル格納
4	音声登録	音声パターンの語番順登録、標準3回学習
5	音声パターンストア	音声登録パターンをファイル単位で格納
6	音声パターンロード	音声登録済みパターンファイルのローディング
7	音声パターン更新	音声登録パターンの更新
8	認識範囲登録	認識テスト時の認識範囲登録
9	認識範囲追加	認識テスト時の認識範囲追加
10	認識範囲解除	認識テスト時の認識範囲解除
11	認識起動	認識準備完了
12	Vモード終了	音声登録テストモード終了

表5中の音声登録テストコマンドを入力することにより、次に述べる音声登録テストモードに移行する。

(2) 音声登録テストモード

音声パターン、表示文字の登録、更新から認識範囲の設定、認識テスト評価の実施までを、音声登録用CRT画面により対話形式で行なうことができる。これら一連の操作に必要なコマンドは、全体で12種類が準備されており、その内容は、表6に示すとおりである。

6 応用例

音声認識装置を利用した音声入力システムには、音声ダイヤル、電話による座席予約、オーダエントリなどの音声入出力システム、計測システム、自動仕分けシステム、検査システム、製品入出庫システムなどがあるが、ここではHR-100IIを用いた音声入力による製品検査業務を紹介する。

製品検査業務には、検査項目ごとの各種点検作業と点検箇

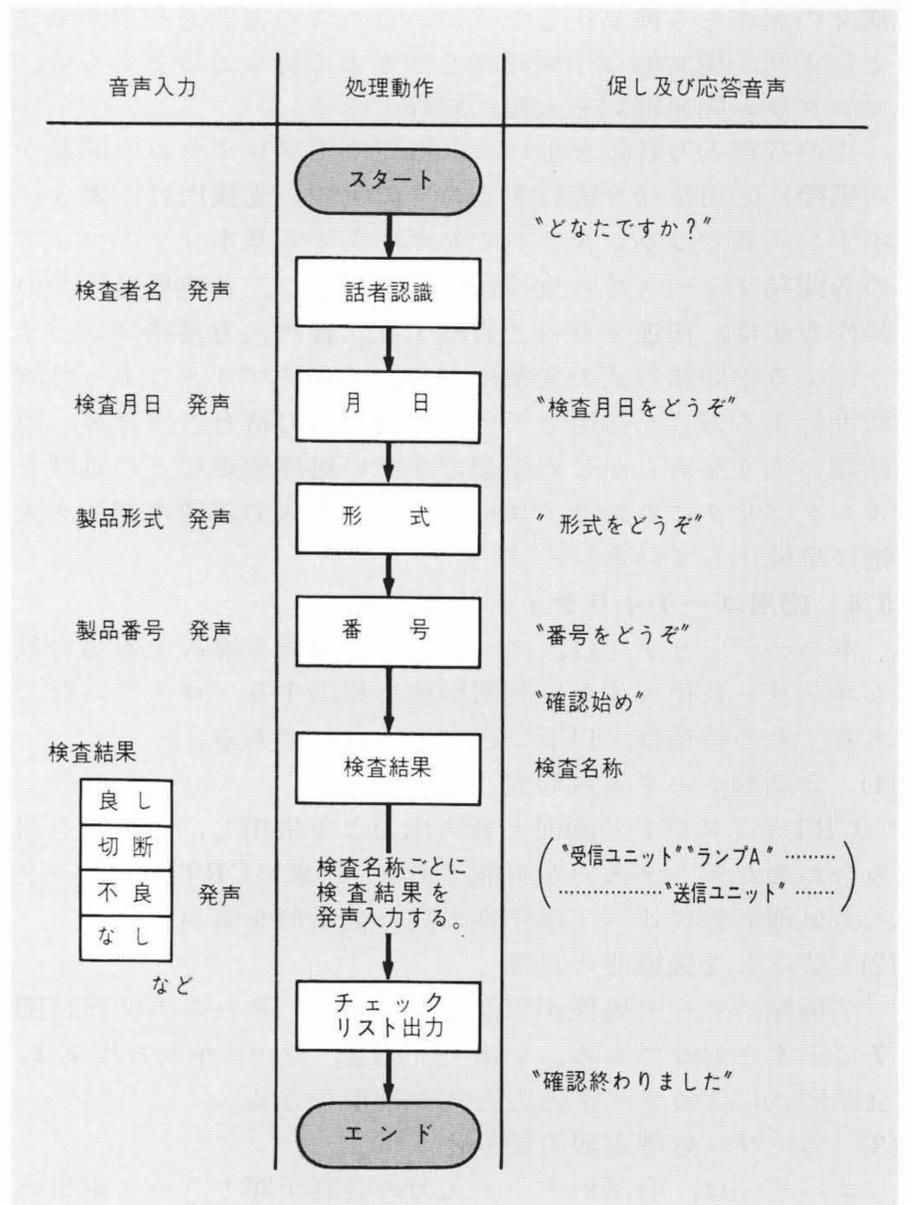


図8 製品検査の検査フロー例 促し音声及び応答音声に応じて検査者が発声すれば、音声認識装置でデータが処理される。

所のデータ測定、及び書類の作成作業がある。このうち、本来人間でなければできない部分だけを人間が作業し、データの作図や作表業務は機械化することにより、検査業務の大幅な省力化が図られる。具体的な検査フローを図8に示す。

HR-100IIからの促し音声により、検査者名、検査月日、検査結果などを発声し機械に入力すれば、検査データの作図、作表は自動的に作成でき印刷出力される。現在これらシステムは、電子装置の組立工場や車両検査工場で既に実用化されている。

7 結 言

以上、音声認識装置とそれを制御し、ユーザーソフトウェア開発を支援するソフトウェアを備えた音声入力システムについて述べたが、今後、音声によるデータ入力手段として、いっそう便利な連続音声認識方式、また、使用する前にあらかじめ音声を登録しなくて済む不特定話者認識方式、更には任意語彙の認識が可能ないわゆる音声タイプライタへの発展が予想されている。これらのニーズに対して、より良い音声入力装置を開発するため、更に努力してゆく考えである。

参考文献

- 1) 姫路, 外: 音声入力応用検査情報処理システム VOCAL の開発と適用, 日立評論, 63, 12, 835~838 (昭56-12)
- 2) 秋本, 外: 音声入力システム用問題向き言語の開発, 日立評論, 63, 12, 831~834 (昭56-12)