

音声入力システム用問題向き言語の開発

Development of Problem Oriented Language for Voice Entry Systems

本稿で述べる音声入力用問題向き言語は、特定話者用の汎用形音声認識装置(HR-100 II¹⁾)の基本的な支援ソフトウェアとして開発したもので、音声入力手順と処理手続きを容易に記述できるようにした音声入力専用的高级言語である。その特長は、複雑な音声入力手順の標準化、木構造形の音声入力構文及び徹底した構造化記述方式の採用にある。

本言語を利用することにより、音声応用業務の音声入力構文の直接記述が可能となり、認識の対象範囲を動的に指定する構文予測認識を自動的に行なえるとともに、十分構造化され保守性に優れた音声応用ソフトウェアを短期間で効率よく開発できる。

秋本文之* Fumiyuki Akimoto

山野紘一* Kôichi Yamano

磯崎真治* Shinji Isozaki

長沢 潔** Kiyoshi Nagasawa

1 緒 言

最近の音声認識とその関連技術の著しい進展につれ、人間にとって最も自然である音声入力がマンマシンシステムの有望な入力手段として注目を集めている。

しかし、実際に音声入力がある業務に適用する場合、高度の音声認識技術をもってしても、音声自体があまりにも人間側に近いがための新たな問題と要請を提起する。例えば、

- (1) 同一環境下で常に安定した発声は望めない。
- (2) 誤発声は日常的で、かつ記録性に乏しい。
- (3) 会話形で実時間応答処理が必要である。
- (4) 非手続き形の実行処理要求が望まれる。

などであり、これら対策上同一時点で認識すべき単語をできる限り少なくすること、認識結果の確認、入力語の取消しのような実行時プロセスの変更などは実用上必須となる。

また、音声入力は第一線の作業現場での直接的な使用形態が多く、入力手順や処理手続きの変更頻度が高い環境にあるのに対し、専門プログラマがいるケースは少ない。

本稿の音声入力言語は、上記のような問題解決を目的に開発したエンドユーザー向け的高级言語である。

2 音声入力言語の特長

本言語は、音声入力応用業務に幅広く適用できることをねらいとした問題向き構造化記述言語である。入力したい単語の集合に対しユーザーが自由に入力順序を示す構文を定義し、入力語が認識されたとき行なうべき操作(アクション)を容易に記述可能としている。主な特長点を以下に示す。

(1) 入力語群の定義

話者が入力する語の集まりを語群化して、構造的に定義できる。また、音声入力実行時に、認識語名と実際の語番との対応づけができる。

(2) 音声入力構文の簡易記述

音声応用業務で使用する入力語の構文図に即して、容易に記述できる。

(3) 標準的な音声入力実行時処理群

音声入力実行時に共通的な処理(認識範囲の設定、確認、取消しなど)を標準化して、システム機能として具備している。

(4) 構造化記述^{*1)}言語

認識語、構文、手続きなどの各定義部を独立させるとともに、入力構文の分岐を制限し簡明な記述形式としている。

(5) ユーザープログラムとの結合

ユーザーが別途作成したプログラムと容易に結合できるように工夫している。

3 言語設計の基本的考え方

3.1 音声入力処理の標準化

音声入力制御の際、ある単語を認識処理するために必要となる標準手順を図1に示す。同図では、同一認識範囲内の単語列(例えば、ある意味をもつ数字列)を入力する場合は特定回数の繰返しで実現することも表わしている。

本言語では、ある入力時点で同一認識範囲に属する語の集まりに対して語群(カテゴリー)を前もって定義し、図1に示した標準入力手順と対応づけて一つの構文としている。換言すると一つの構文は、認識語、構文分岐、前処理、後処理、反復回数などから成ることを意味する。ある音声応用業務はこの構文を単位とし、組合せ順序づけにより全体の入力構文を形成する。

したがって、ユーザーにとっては、図1の音声入力手順と上述の入力構文をどれだけ簡単に記述できるかが大きな関心事となる。このため、本言語システムの標準的な支援機能として、認識範囲の自動設定、促し、復唱、取消しなど処理手続き群を具備させ、音声入力にかかわるユーザー負担の最小化を図った。

3.2 音声入力構文

音声入力の代表的な応用例として、荷物の番号、経路、行先地を音声入力することによって仕分け制御する簡単なモデルの入力構文図を図2に示す。同図中の円内のS0、S1などの記号は、一つの構文の代表名であり、矢印は次に入力される語を規定する構文への状態移動を示す。本構文図を前もって登録しておくことにより、入力語の予測認識が可能となる。

※1) プログラム構造の簡明化を目的とした手法で、ソフトウェアの開発・保守が容易になる。

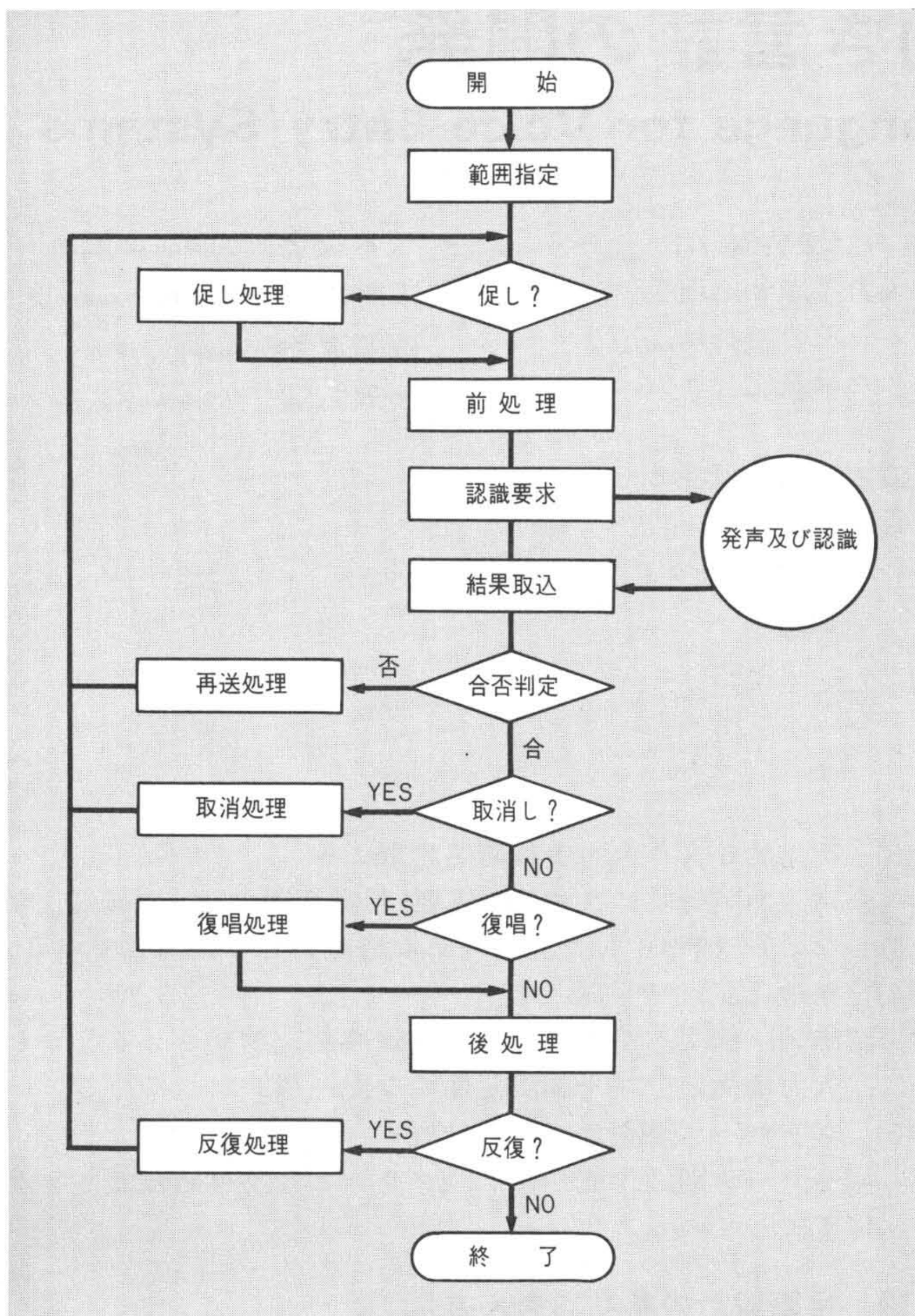


図1 音声入力処理の標準手順 同一認識範囲内の単語又は単語列を、入力処理するために必要な標準手続きを示している。

一般的にこのような構文を定義するためには、同図中にも表現されているように構文の枝分かれ、逆戻り、反復の記述が必要となる。

これらの記述を無秩序に許すと、構文は複雑になりすぎて解析が困難となる。このため、入力構文の分岐は親子関係を表現する木構造形式を基本とし、直系上位及び下位への分岐だけを認め無制限分岐を禁止した。

3.3 言語システムの構成

音声入力がある業務に適用する場合、まず図2に示したような構文図が必要であり、これをもとに音声入力プログラムを開発することになる。本言語プログラムは、HR-100IIシステム自体で開発が可能であり、同じくマイクロコンピュータシステム開発装置SD20²⁾でも可能である。

ある音声応用業務システムを開発する場合の開発支援ソフトウェアの関係をまとめ、音声入力言語で記述した原始プログラムは、画面エディタを用いてフロッピーディスクファイルに格納し、次に音声入力言語コンパイラによって、いったん中間語に翻訳し、制御プログラムのもとでインタプリタによって解釈実行する。

3.4 音声入力処理の実行制御方式

実際の音声入力処理の実行時に問題となる点は、異常事態の発生への対処方法である。機器の誤動作のほか、発声すべき音声の間違った場合や不明瞭な発声で誤認識した場合の措置などは、音声入力実行時に随所で要求される機能であり、

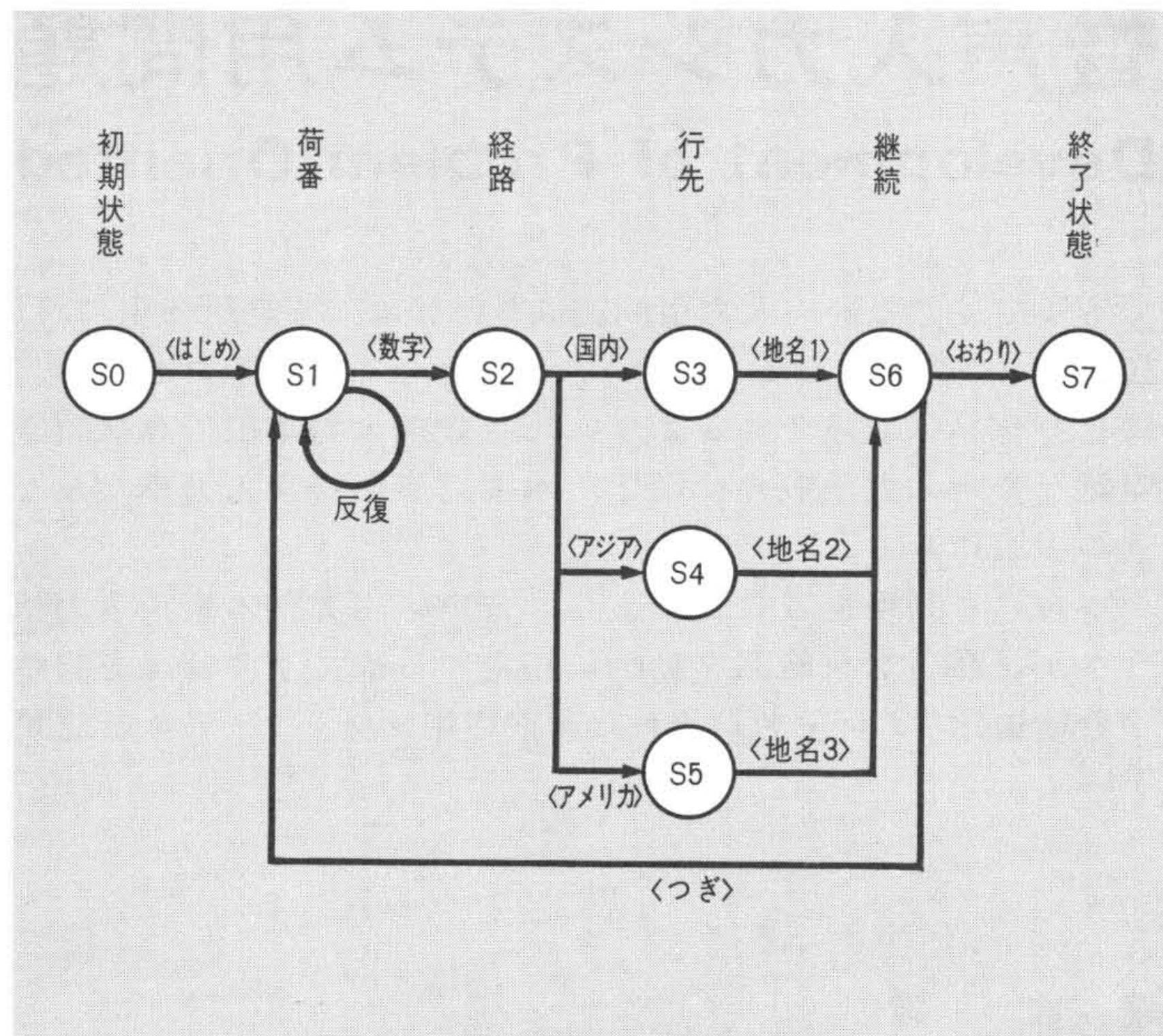


図2 荷物仕分け制御の音声入力構文図 荷物の番号、経路、行先地などを音声入力することによって、仕分け制御する場合の入力構文を示す。

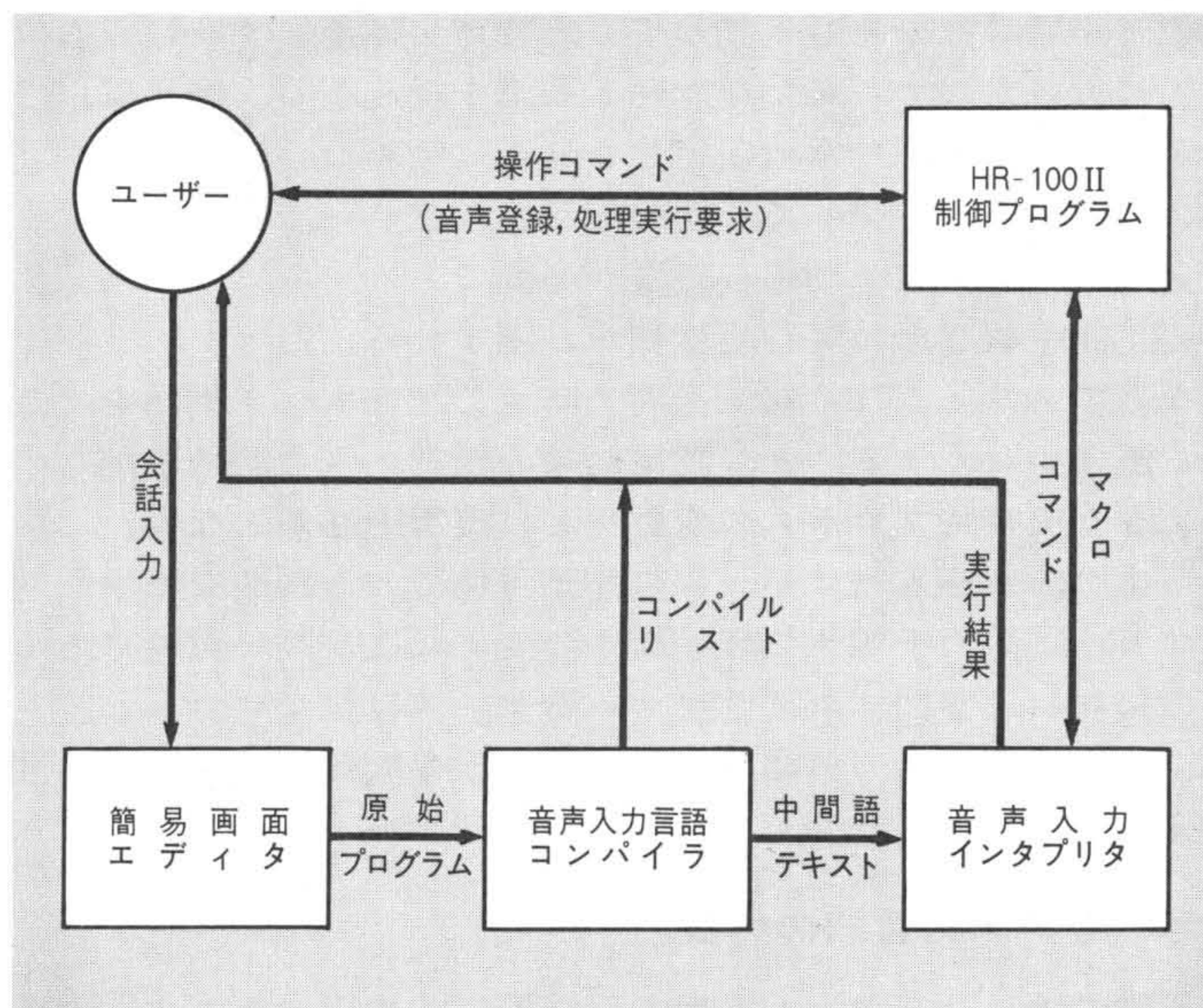


図3 音声入力言語システムの全体構成 音声入力装置(HR-100II)を用いて、音声入力応用プログラムを開発する場合の支援ソフトウェアの関係を表わしている。

またそれらは使いやすくするために動的に変更できることが望まれる。

本システムでは、これらの要求にこたえるため、原始プログラムをコンパイラでいったん中間語テキストに変換し、インタプリタによって実時間で解釈実行していく方式を採用し、異常処理は実行時に呼び出される標準ルーチン(実行時ルーチン)で対処できるようにした。前掲図1の音声入力標準手順に合わせて、実行時ルーチンとしては次の6種類を準備した。

- (1) 取消し処理〔CANCEL〕
- (2) 音声認識確度〔SCORE〕
- (3) 認識レベル差〔LEVEL〕
- (4) 促し処理〔PROMPT〕
- (5) 復唱処理〔ECHO〕

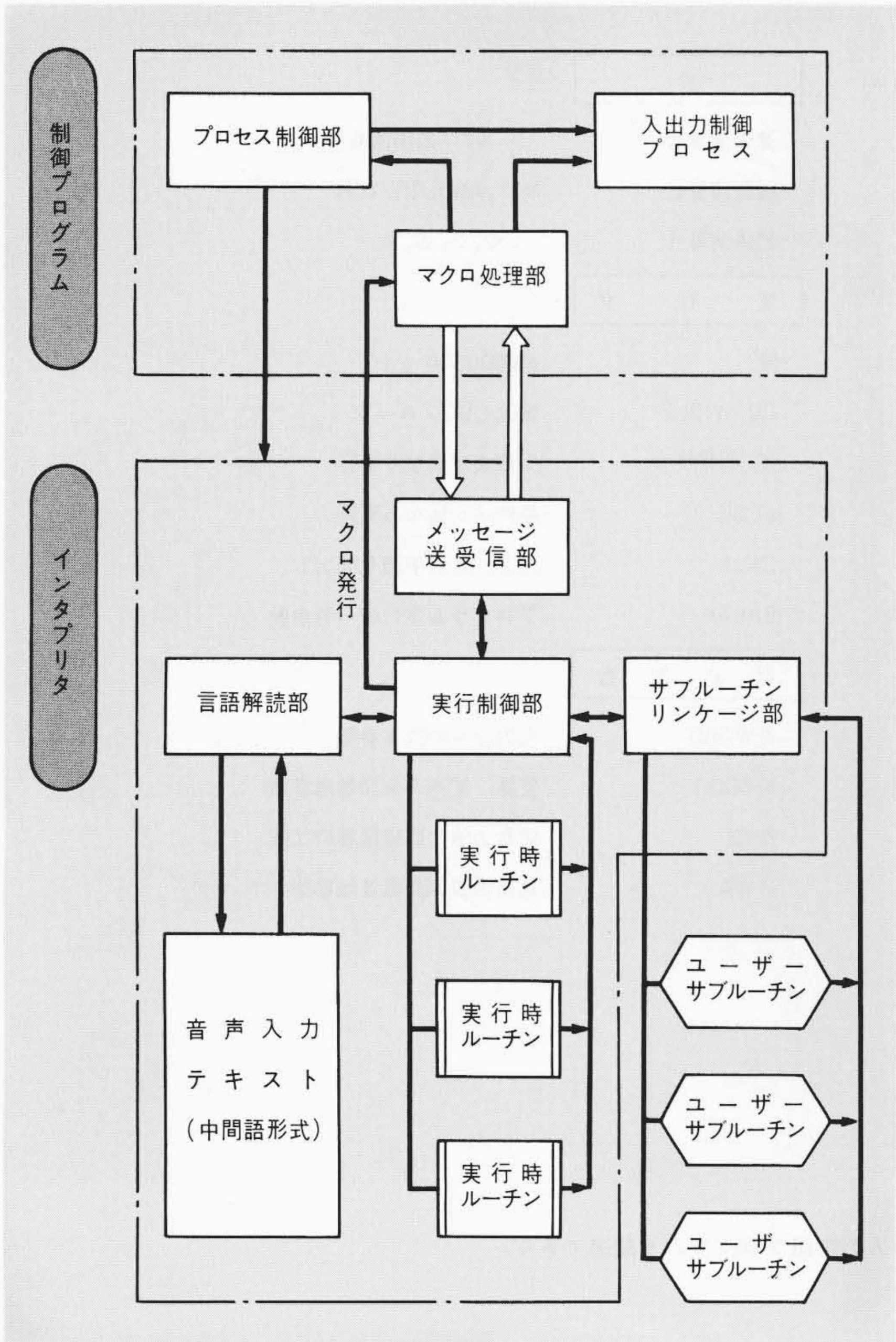


図4 音声入力処理の実行制御方式 インタプリタの処理構造と制御プログラムの関係を表わしたもので、コンパイル結果後の中間語形式プログラムの解釈実行方式を示す。

(6) デバッグ[DEBUG]

これら実行時ルーチンは、各種指定変数の記述をも含めて言語仕様に織り込むこととした。

インタプリタの主要機能を要約すると、

- (1) 中間語形式の音声入力テキストの解釈実行
 - (2) 制御プログラムとの送受信マクロ処理
 - (3) 実行時ルーチンの起動制御
 - (4) ユーザーサブルーチンへの分岐制御
- などである。

インタプリタと制御プログラムとの関係を、音声入力処理の実行制御方式としてまとめて図4に示す。

4 言語仕様

4.1 プログラムの構成

一つの音声入力処理プログラムの構成を図5に示す。音声入力条件、認識単語、処理手続き、構文定義部など、全体は六つの独立した機能で構成される。

音声応用業務処理の実行で最も大切な構文定義部の記述例を図6に示す。同図は、前掲図2の荷物仕分け制御の入力構文図を音声入力言語で記述したもので、簡明な表現になっている。なおこの例では、音声入力に関する標準処理をすべて利用するものとして、前処理手続き名、後処理手続き名の記

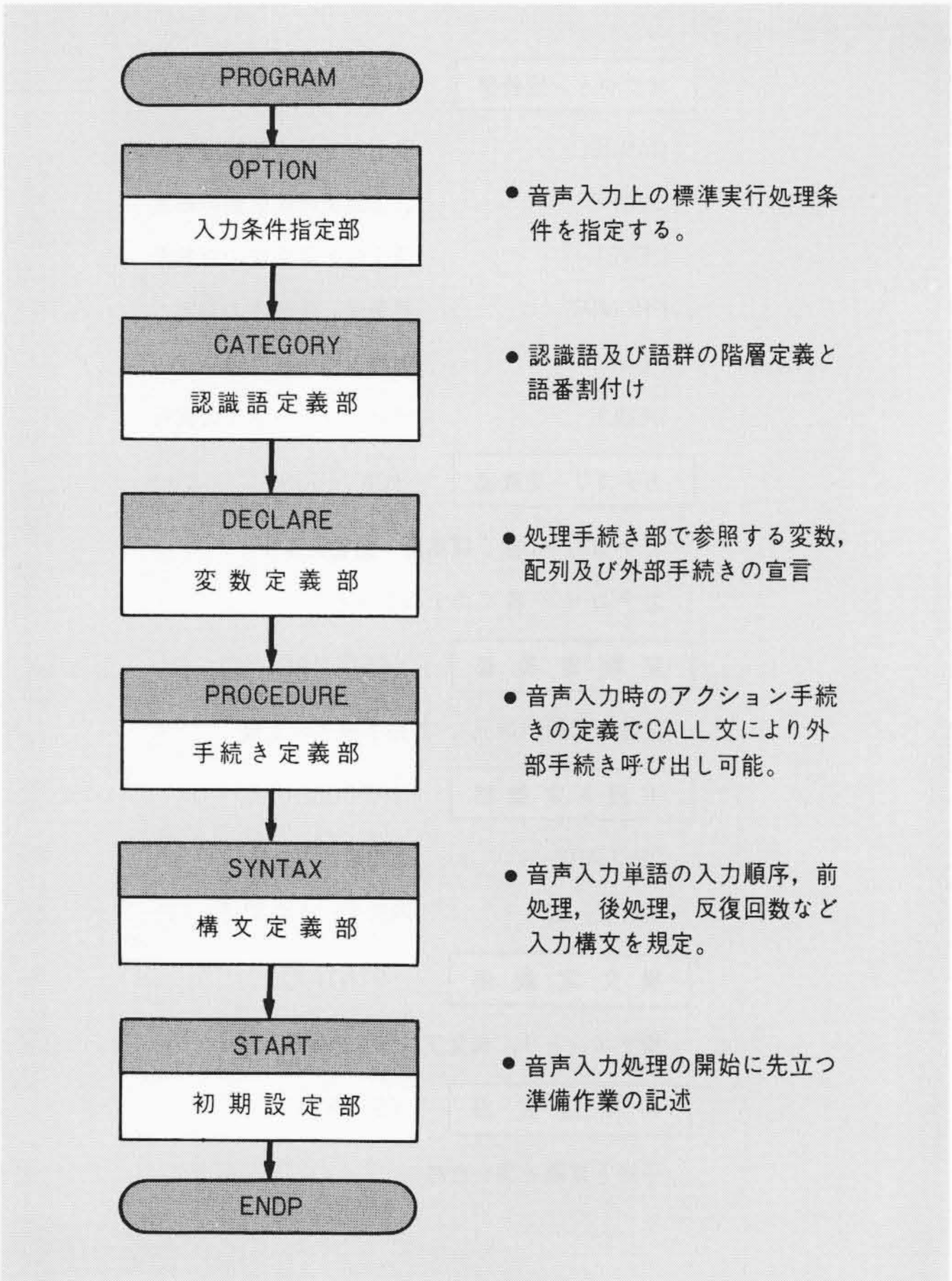


図5 音声入力応用プログラムの構成 本図に示す各部を音声入力言語で記述することによって、音声応用業務システムが開発できる。

00053:	/*	SYNTAX	DEFINITION	PART	*/				
00054:		S0:SYNTAX							
00055:		1:(HAJIME	->	S1)					
00056:		ENDS							
00057:		S1:SYNTAX							
00058:		1:%NUMBER			/5				
00059:		2:(KOKUNAI	->	S3,					
00060:		ASIA	->	S4,					
00061:		AMERICA	->	S5)					
00062:		3:(TSUGI	->	1,					
00063:		OWARI	/	PEND)					
00064:		S3:SYNTAX							
00065:		1:%CHIMEI1							
00066:		ENDS							
00067:		S4:SYNTAX							
00068:		1:%CHIMEI2							
00069:		ENDS							
00070:		S5:SYNTAX							
00071:		%CHIMEI3							
00072:		ENDS							
00073:		ENDS							

図6 音声入力構文の記述例 前掲図2の仕分け制御モデルの音声入力構文部分を、音声入力言語で記述した例である。

オプション定義部 (OPTION) CANCEL 取消しを示す認識語の指定 SCORE 許容認識確度点数の指定 LEVEL 第1位と第2位の点数差 PROMPT 標準促し手続きの指定 ECHO 標準復唱手続きの指定 DEBUG デバッグモードでの実行	カテゴリ定義部 (CATEGORY) カテゴリー名：認識語＝語番又は＊ カテゴリー名：カテゴリー名	変数定義部 (DECLARE) 変数，配列(1次元)，外部手続きの定義	手続き定義部 (PROCEDURE) DECLARE 局部変数の定義 実行文，式 内部手続きの記述	構文定義部 (SYNTAX) 構文エントリ，構文ブロックの定義	初期設定部 (START) 手続き定義と同一仕様	式 算術演算式 +-*/ (四則演算子) 論理演算式 NOT, AND, OR, XOR 関係演算式 =, <, >, ≤, ≥	実行文 DO 単純DOブロック DO WHILE 繰返しDOブロック IF THEN 条件式が真なら実行 ELSE 条件式が偽なら実行 CALL 内部，外部手続き呼出し BREAK プログラム実行の一時中断	組込関数 @WORD 入力音声の語番参照 @ADDR 変数，配列の先頭番地参照 @ID 変数の値と認識語番の比較 @WA 認識語表の先頭番地参照
---	---	--	---	--	---	--	--	---

図7 音声入力言語の詳細仕様 本図に示した言語の使い方を知れば，音声入力応用プログラムを記述できる。

述を省略している。
4.2 言語仕様の詳細

本言語の詳細仕様を図7に示す。同図では図4に示したプログラム構成の各定義部にほぼ対応している。式，実行文，組込関数は，手続き定義部及び初期設定部で記述可能である。ユーザーサブルーチンなどの外部手続きを呼び出す場合は手続き定義部でCALL文を用いる。呼び出す外部手続きは，変数定義部で宣言し，引き数(最大10個)は手続き部で宣言することによって簡単に引き数の受渡しができる。実行文などの一般的な記述形式はPL/H³⁾スタイルとし，ユーザーが新たに覚える部分を少なくするとともに，無条件Jump命令を禁止するなど，言語仕様上から必然的に構造化されるようにしている。

5 適用結果と検討

本言語は日立製作所社内の数事例システムに適用し，その有効性が確認されている。今後更に多くの応用が計画されている。本言語システム採用の第一の利点は，音声入力構文の記述性にある。これをアセンブリ言語で記述する場合の労力に比べ，開発期間で約半に短縮できる。この比較では，本言語システムがもつ標準入力処理手続き群などの開発工数は除外している。第二の利点は，ユーザープログラムとの結合の容易性である。このため，既に稼動している生産合理化システム(例えば，部品加工，自動倉庫，検査業務など)に音声入力を比較的容易に導入できた⁴⁾。第三の利点は，徹底した構造化記述方式としたため，プロ

グラムの共同開発が容易になりデバッグ面，保守面でも顕著な効果が認められた。今後，標準実行時処理群などを拡充し，よりいっそう記述性を向上していく必要がある。

6 結 言

音声入力応用業務プログラムの開発負担の軽減を目的として，音声入力言語及び言語プロセッサを開発した。本言語システムは，汎用形音声入力装置(HR-100II)で稼動するもので，音声入力構文を容易に記述でき，しかも自動的な構文予測認識を可能としている。また，音声入力にかかわる複雑な処理手順は，すべてシステム標準機能とし，かつユーザー固有処理の記述も可能とした。本言語の仕様設計では，プログラム構造の簡明化に多くの工夫を施した。このため，プログラム開発，保守性に優れた能力をもっている。今後本言語が広く利用されるように努力を重ねる予定である。

参考文献

1) 木村，外：汎用形音声入力装置“HR-100II”の開発，日立評論，63，11，825～830(昭56-11)
2) 日立製作所：日立マイクロコンピュータシステムFDOSユーザーズマニュアル(昭53-2)
3) 吉村，外：マイクロコンピュータ用高級言語PL/Hシステム，日立評論，61，4，305～310(昭54-4)
4) 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，63，11，835～838(昭56-11)