

視覚情報からの自然語文の生成について

高木 朗 伊東幸宏 北國和憲 清水正朗 小原啓義

(早稲田大学 理工学部)

1. はじめに

自然言語処理の応用範囲を広げる為には、視覚情報と言語情報の係わり合いについて検討することが重要である。しかしながら、この点について考察した研究は、二三の例[1][2]を除き、ほとんど見受けられない。我々は、視覚情報と言語情報の対応関係について検討を加え、視覚情報の特徴解析結果から自然語文を生成するシステムを試作した。本稿では、このシステムの構成について述べる。

2. システムの概略

図1にシステムの概略を示す。システムに与えられる入力は、図形処理[3]の結果得られる特徴テーブル(図2)である。このテーブルは、画面中の図形の形状に関する特徴をまとめた名詞特徴テーブル(noun feature table, 以下、N.F.T.)と、動詞情報をまとめた動詞特徴テーブル(verb feature table, 以下V.F.T.)、図形の位置を表すLOCATION、動詞の相情報を表すASPECTから成る。これらをまとめて命題特徴テーブル(proposition feature table, 以下P.F.T.)と呼ぶ。尚、静止画の場合はN.F.T.のみが与

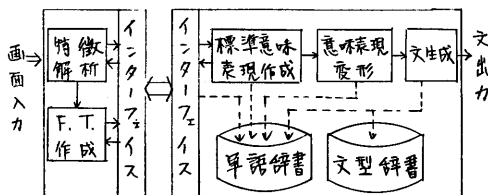


図1 システムの概略

(OBJECT 1 (SIZE 910))	(COLOR (HUE 15))	(SUBPART 1 (SIZE 231))	(COLOR (HUE 15))	(SHAPE (TRIANGLE))	(SUBPART 2 (SIZE 110))	(LOC (SUBPART 1) (X AXIS (+))))
(VFT (OBJECT 2 (SIZE (CHANGE (CONSTANTLY) (RATE 8))))						
(C) (LOCATION (((ZAXIS (+)) ((OBJECT 3))))						
(D) (ASPECT (PROGRESSIVE))						

命題特徴テーブル(P.F.T.)

図2 特徴テーブル

えられる。N.F.T.は図形毎に作られ、表1(a)に示す特徴が書かれている。V.F.T.は属性値変化がある図形又は部分図形の属性毎に作られ、表1(b)の特徴が書かれている。又、LOCATIONは、着目している図形の最近傍の図形に対する相対位置を表し、ASPECTは、属性値変化が開始したか、継続しているか、終了したかを表す。P.F.T.は一つの図形毎に作られる。同一図形(部分図形)の複数の属性が変化して

図形 (部分 内部図形)	面積 色 形 長さ 巾 指示情報 部分1 : 部分n 内部図形1 : 内部図形m	部分	面積 色 形 長さ 巾 指示情報 主要部分との位置関係	値 色相 値 値 値 値 値
図形 (内部図形)	部分 属性	一定の変化 方向 周期的な変化 割合 方向 振幅 周波数		

表1 検出する特徴

(A) 名詞

- a. 図形を表す名詞 「ロボット」「木」等
- b. 部分図形を表す名詞 (以下、部分名詞) 「コップの取手」等
- c. 属性を表す名詞 (以下、属性名詞) 「色」「数」「位置」「速さ」等
- d. 数詞 「1本」「3つ」等
- e. 代名詞 「これ」「これら」

(B) 動詞

- a. 属性値の時間的変化を表す自動詞 「動く」「転がる」等
- b. 「もっている」「だ」「である」「ある」「等しい」
- c. 形容詞、形容動詞 終止形 (原級、比較級、最上級) 「赤い」「大きい」「～より大きい」「～の中で一番大きい」等

(C) 述修飾表現

- a. 属性値を言及するタイプ
 - a1. 属性形容詞連体形
 - i) 値を直接言及するもの 「赤い」「丸い」等
 - ii) 標準値との比較結果を言及するもの 「大きい」「広い」等
 - a2. 属性値を言及するタイプの形容詞に類した表現
 - i) 「～の」 「10の面積の」等
 - ii) 所有格関係節 「面積が10の」等
 - iii) 主格関係節 「10の面積をもつ」等
- b. どの様な部分をもっているかを言及するタイプ
 - i) 主格関係節 「長い足をもつテーブル」等
 - ii) 所有格関係節 「足が長いテーブル」等
- c. 数を言及するタイプ 「3つの」「5本の」等
- d. 部分名詞を修飾するタイプで、その部分名詞がどの図形の部分であるかを言及するもの 「このコップの取手」等
- e. 属性名詞を修飾するタイプで、どの図形の属性が問題となっているかを言及するもの 「ボールの色」等
- f. 指示を表すタイプ 「この」「これらの」

(D) 連用修飾表現

- a. 属性値を言及するタイプ
 - a1. 属性形容詞連用形 「速く」等
 - a2. 副詞 「ゆっくり」等
 - a3. 副詞に類した表現 「～の速さで」「～に向かって」等
- b. 位置を言及するタイプ 「机の上を」「家の前を」等
- c. 同時に起っている他の属性値変化を言及するタイプ 「転がらみながら」「膨らみながら」等

表2 取り扱う語句

値をもたない時、その値は意味表現中から削除されている。
 そうした属性毎にdefault value が記憶されている。

属性値変化を表す自動詞（「膨らむ」、「動く」等）の意味表現は図4の様に表される。即ち、表1(b)の特徴を「～の属性に関して～の割合をもって～の方向に向かって変る」という形で意味表現化した構造をもつ。但し、「方向」は属性名が位置の時には進行方向を、それ以外の属性の時には属性値変化の微係数の符号を表現する。属性値が不定の場合には、その属性に関する枝が削除され、default value が記憶されている。尚、同一図形の複数の部分の属性値変化あるいは同一図形の複数の属性の変化が複合して一つの動詞の意味が構成される場合（「転がる」、「歩く」等）、図4の意味表現が部分毎、属性毎に作られ、それらが一つにまとめられて動詞全体の意味表現が構成される。但し、その際部分の変化が複合している場合には主要部分の属性値変化を主動詞に、又、複数の属性における変化が複合している場合には、その内の外部から指定された属性の変化を主動詞にして、他の属性値変化の意味表現を「～しながら」、「～することによって」等を介して主動詞に接続する形をとっている。更に、「歩く」の様に生物主語をとる動詞は、各部分の属性値変化の意味表現に、「～と～を」という表現を付加している。

3.3 意味表現の変形

本システムでは、同義な言い換え表現を生成する為に、意味表現の変形処理を行う。図5に代表的な意味表現変形規則を挙げる。

以上述べてきた意味表現は次の様な性質をもつ。

- ① 単語の意味として記憶されている意味表現を、意味的係り受けに従って接続することにより、句、節、文の意味表現を構成することが出来る。
- ② 逆に意味表現が与えられると、単語辞書を参照しつつ、それを単語毎の意味表現が接続し合った形に分割することが出来る。従ってそこから意味表現を表層文に変換する際に必要な単語の組、及び、それらの単語の意味接続関係を知ることが出来る。
- ③ 一つの意味表現は変形処理を行うことにより、同義な表層文を生成する為の他の形の意味表現に変形することが出来る（但し、任意に変形出来る訳ではない）。

本システムでは②、③の性質を利用している。

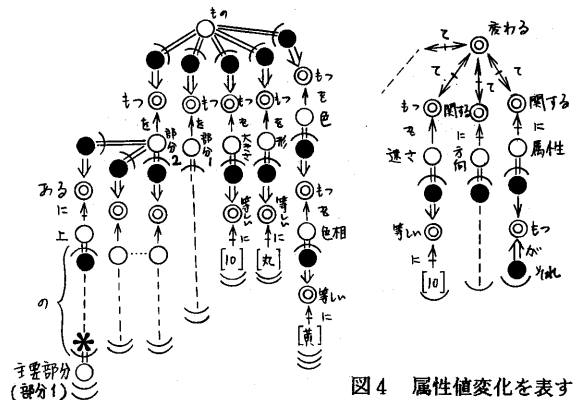


図4 属性値変化を表す

図3 図形名詞の意味構造

自動詞の意味構造

(I) 主節、従属節の反転 (a) 赤いボールが机の上にある (b) 机の上にあるボールは赤い 机の上にあるボール 赤い (a) ○のキ→○ ←○ キ赤い) 赤い ボール ある 上 の 机 (b) (赤いキ回← ○キ●→○←○のキ○)	(III) 「AのBのCが～」→「AのBはCに関して」の変形 (a) ロボットの手の位置が変っている (b) ロボットの手は位置に関して変っている が (a) いる○←○のキ○のキ○ て位置 手 ロボット 変る○ は手 (b) いる○←○のキ○ロボット て 変る○ ←○ →○ ←○ て 関する に 位置	(V) that clause 中の主格に関する変形 (a) ロボットは手が動くということを引き起している (b) ロボットは手に対してそれが動くということを引き起している ロボット いる 引き起す こと 動く 手 は て を という が (a) ○ → ○ ←○ → ←○キ● ←○ ←○) ロボット いる 引き起す こと 動く それ は て を という が (b) ○ → ○ ←○ → ←○キ● ←○ ←○) て 対する に
(II) 関係節構造の変形 (a) 赤いボール (b) 赤い色をしたボール (c) 赤い色のボール (d) 色が赤いボール ボール 赤い: X (a) ○キ●→○ ←○キ●→○ ←○キ●→○ ←○ (赤)) ボール した 色 赤い: Y (b) ○キ●→○ ←○キ●→○ ←○キ●→○ ←○ (赤)) ボール の 色 赤い (c) ○キ●→○ ←○キ● → Y ... ボール 色 赤い (d) ○キ●キ● → ○ ←○キ● → 回○キ● ... Y ...)	(IV) 「～は～をもつ」→「～には～がある」の変形 (a) コップはとってをもつ (b) コップにはとってがある コップ は もつ を とって (a) ○ → ○ ←○ ←○ コップ には ある が とって (b) ○ → ○ ←○ は副動詞「は」を表す。	(VI) that clause のto不定詞変換 (a) 位置が変るということを引き起す (b) 位置に関して変るということを引き起す (c) うごかう 引き起す こと 変る 位置 を という が (a) ○ ←○ ←○ ←○ ←○ 引き起す こと 変る 関する 位置 を という が (b) ○ ←○ ←○ ←○ ←○ す 動く 引き起す こと 変る 関する 位置 を て に (c) ○ ←○ ←○ ←○ ←○

図5 意味表現の変形の一例

4. 処理手順

視覚情報から言語情報を生成する際、一般に同一画面から生成し得る文は一意には定まらない。例えば、机の上にコップが描かれている画面の場合、「このコップは赤い」「机の上にコップがある」等を始めとして、生成し得る文は極めて多数である。従って、画面から生成する文の数を制限でき、しかも要求があれば様々な文が生成できることが望ましい。そこで、本システムでは以下の手順で処理を行っている。図6に処理手順の概略を示す。

4.1 テーマの決定

一つの画面から生成される文のタイプは、画面中のどの図形に着目して、何をテーマとするか（その図形のどの性質について言及するか）によって大別される。そこで本システムでは、画面中に一つしか図形がなければその図形に着目し、又、複数の図形が存在する場合には、言及する図形を選択する為の指示を表す矢印を画面中に置いて着目する図形を決定している。画面中で指示を受けている図形は、

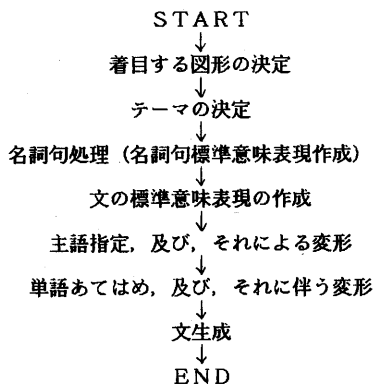


図6 処理手順の概略

N.F.T.中にその旨が記録されている。又、テーマは表4に示すものを取り扱う。

4.2 標準となる意味表現の作成

着目する図形及びテーマが決定された後も、表4に示した様に生成され得る文は依然として多様である。本システムでは、画面の解析結果からまず標準となる意味表現（以下、標準意味表現）を作成し、これに前述の変形を施して様々な文を表す意味表現を作るという方針をとっている。この標準意味表現は動画の場合にはP.F.T.から作られるが静止画の場合には、例えば比較情報の様に画面の解析結果中には陽には存在しない情報もあり、テーマ毎に個別な処理を必要とする。但し、いずれの場合も画面中の図形を表す意味表現は共通にN.F.T.から作られる。そこで、以下に図形を表す名詞句意味表現と各テーマ毎の標準意味表現の作成手順を示す。

4.2.1 名詞句の処理

視覚情報中の図形は、その中に名詞情報と連体修飾情報とが混然一体となって含まれている。従って図形の解析結果であるN.F.T.から直接作られる意味表現（これを名詞句原始表現と呼ぶ）中にも、名詞情報と連体修飾情報とが一体となって含まれている（全く同様な事が動詞情報についても言える）。ここで行う名詞句の処理は、この原始表現から両者を分離することを第一の目的としている。

以下に処理手順を示す。

- i) 名詞句原始表現の作成：N.F.T.の各枝を「～を内包するところの」「～に等しいところの」等の形で関係節化し、図3の形の意味表現を作る（図7(1)）。
- ii) 概念化：i) で作られた名詞句原始表現を単語辞書中の名詞意味表現と比較し、着目している図形が属する概

表4 テーマの指定

画面	テーマ	外部からの指定モード	生成し得る文例
静止画	着目した図形が属する概念のクラスについて	INC	「これはコップだ。」 「これらはコップだ。」
	着目した図形の存在、又は位置について	EXIST	「赤いコップがある。」 「赤い色のコップがある。」 「色が赤いコップがある。」 「大きな机の上に赤いコップがある。」
	着目した図形がもつ図形的特徴（色、大きさ、形等）について	POSS (ATTRIBUTE) 着目する特徴名	「このコップは赤い色をしている。」 「このコップは赤い。」 「このコップの色は赤い。」 「このコップはあのコップより大きい。」 「このコップは三つのコップの中で一番大きい。」
	着目した図形がもつ部分図形について	POSS (SUBPART)	「このコップは赤い取手をもっている。」 「このコップには赤い取手がある。」 「このコップの取手は赤い。」
動画	着目した図形の属性値の変化について	—————	「ボールの位置が木に向かって変っている。」 「ボールは回転しながら木に向かって動いている。」 「ボールは木に向かって転がっている。」

念を決定する(図7(1))。その際、図中に示した条件を用いる。これは通常、人間が文中で最も抽象度の低い(最も具体的な)名詞を用いる傾向があることによる。

iii) 連体修飾表現の切り出し: 名詞句原始表現の、名詞概念からはみ出した枝(図7(1) b, c, d)を根本まで遡って切り出す。これは原始表現に於てはめられた名詞概念がその枝に対応する属性において一定の属性値をもたないことを意味し、この部分が名詞に掛る連体修飾の部分となる。名詞概念を表すatomに連体修飾表現を接続した形(図7(2))を名詞句標準意味表現とする。

iv) 連体修飾表現のタイプ指定: 名詞句標準意味表現中の連体修飾表現の各枝について、表5の形式で表層における文型を指定する。これにより連体修飾表現の意味表現を変形する(図7(3))。表中の「標準値との比較結果を言及する」とは表層上で属性値を直接言及せず、default value との比較結果を+/-で表すことを言う(この場合、この枝は後に形容詞化される)。

4.2.2 各テーマ毎の標準意味表現

ここではテーマ別に動詞まわりの意味表現を作り、名詞句処理の結果とを合せて文の意味表現を作る。

(1) 「着目した図形が属する概念のクラス」を指定した場合 この場合は着目した図形がどのような概念に属するかを

4.2.1 i) ii)と同様な方法で調べ、その概念を表すatomを用いて表6 2-Aの形の意味表現を作成する。

(2) 「着目した図形の存在または位置」を指定した場合 画面中に着目した図形以外の図形が存在しない時には、着目した図形の存在を表す表6 2-B i)の形の意味表現を作る。他に図形が存在する場合は、着目した図形に最も近い図形との相対位置を求め、表6 2-B ii)の形の意味表現を作る。

(3) 「着目した図形がもつ図形的特徴」を指定した場合

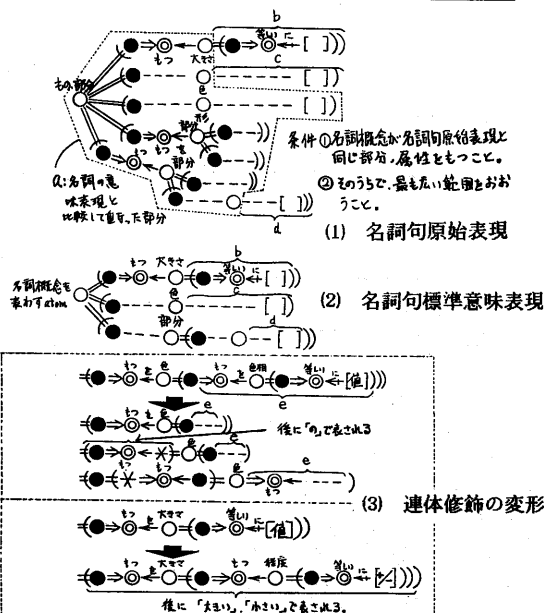


図7 名詞句処理

この場合、着目した図形のどの特徴を言及するかについても同時に外部から指定される。ここで作られる意味表現は、

①着目した図形と同じ概念のクラスに属する図形が他にない場合(例えば、コップに着目した場合、他にコップがない場合)、②他に一つある場合、③他に二つ以上ある場合、で若干構造が異なる。①の場合、着目した名詞句原始表現から指定された特徴の記述を抜き出し、表6 2-C i)の形の意味表現を作る。②③の場合は、着目した図形と同じ概念のクラスに属する図形に対して名詞句処理を施し、その結果得られる名詞句原始表現、及び着目した図形の名詞句原始表現から、外部より指定された特徴の属性値を比較することにより表6 2-C ii) iii)の形の意味表現を作る。これらの意味表現は、原級、比較級、最上級の意味表現となり、後に必要に応じて形容(動)詞化される。

(4) 「着目した図形が持つ部分」を指定した場合

この場合、着目した図形のどの部分を言及するかが、やはり画面上で矢印によって指定されている。これに従って着目している図形の名詞句原始表現から該当する部分図形の記述を抜き出し、名詞句処理を施す。その結果を用いて表6 2-Dの形の意味表現が作られる。

(5) 画面中に属性値変化がある場合

この場合動詞まわりの意味表現はP.F.T.を直接変換することにより作られる。手順は以下の様になる。

- i) 着目した図形を含むP.F.T.を探す。
- ii) そのP.F.T.中のN.F.T.部に記された図形に対し名詞句処理を行う。
- iii) V.F.T.を意味表現に変換する。V.F.T.は動詞のprimitive「変る」を用いて表現されている為、ここで変換された意味表現は属性を主語にして「～の図形の(～の部分)属性が変る」という形となる。副詞情報は「～の割合をもって」「～の方向に向かって」等の形で意味表現化される(図4)。その際、V.F.T.中にあった図形名詞部分名詞は名詞句処理の結果に置き換えられる。同一図形の複数の部分の属性値変化、もしくは同一図形の複数の属性の変化が存在する場合には、各属性値変化を表すV.F.T.に対して上述の処理を行い、主要部分(前述)に

		パラメータ
表層に出力しない	値を直接言及する	0
	値を直接言及する	1
表層に出力する	主格関係節で表す	2
	所有格関係節で表す	4
	格助詞「の」を用いて表す	3
	標準値との比較結果を言及する	10

表5 連体修飾の表現形式の指定

関するV.F.T.又は変化している複数の属性のうち主要な属性（外部から指定される）に関するV.F.T.から得られた意味表現を主動詞として、これに他のV.F.T.から作られた意味表現を「～ということに伴って」、「～ということによって」等の意味表現を介して接続させる。
 iv) LOCATIONの内容を意味表現に変換する。
 v) ASPECTの内容を意味表現に変換する。
 vi) 以上をまとめて文の意味表現とする（表6 2-E）。

4.3 主語の指定とそれに伴う変形処理

標準意味表現が作成されると、そこから同義文を生成する為に意味表現の変形を行う。この為、まず生成したい文の主語となる単語を外部から指定し、これを手掛りにして変形を行う（標準意味表現はテーマ毎に決った構造をもち、指定された語を主語にする為の変形は予めわかっている。

現在の所、出力文数を制限する為、表6に示す範囲の名詞に対してのみ、変形処理を行っているが、これを意味表現全域に拡張することは容易である。尚、属性値変化をテーマとする時に生物主語が指定されると、3.2で述べた「～ということを引き起す」という表現を付加するか否かが同時に指定される。処理手順を以下に示す。

- i) 標準意味表現上で主語になり得る名詞（主格名詞及び「もつ」の目的格につく名詞、表6 2参照）を一つ選び主語として指定された語と比較する。
- ii) 両者が一致すると、その位置に固有の変形を施す（表6 3）。又、テーマが属性値変化の場合で、生物主語が指定され、「～ということを引き起す」を付加する様に指定されている時にはこれを付加する。
- iii) 上述の処理を、標準意味表現上の主語となり得る位置にある全ての名詞について繰り返す。

表6 標準意味表現及び主語指定による変形

1 モード	2 標準意味表現	3 主語指定により変形された意味表現	4 用いた変形の種類の
A INC		_____	変形なし
B EXIST		_____	変形なし
C POSS (ATTRIBUTE)			変形なし
			IV I
D POSS (SUBPART)			変形なし
			IV I
			I
E 属性変化の自動詞			変形なし
			IV

4.4 単語あてはめによる変形

この段階までに作られる意味表現は、図形及び部分図形を表す名詞句以外は全てprimitiveで記述されており、表層の単語との対応はとれていない。そこで次にこの意味表現を、表層と対応がついた概念(単語)を表すatomを用いた表現に変換する処理を行う。処理手順を以下に示す。

- i) 処理すべき意味表現を、辞書中の単語意味表現とheadをそろえて比較し、前者が後者を包含するか否かを見る。
- ii) 包含しない場合には、他の単語についてi)の処理を繰り返す。
- iii) 包含した場合、その単語で被われた意味表現上の部分を、その単語を表すatomで置き換える。そこからはみだした意味表現の枝の各々についてi)からの処理を再帰的に実行し、その結果を先程置き換えられた、単語を表すatomに接続させる。

但し、その際、以下の二点に注意する必要がある。

- ① 単語の意味表現と比較する際、処理すべき意味表現を変形して初めて包含関係が成立する場合がある(「の」、「ながら」及び属性値変化を表す自動詞の場合等)。この時はその変形を施した後に単語の意味表現と比較する。
- ② 形容詞、形容動詞終止形は表3に示した形の意味表現で表される。本システムでは回はそれに続く関係節、名詞節中の動詞を引き出して主動詞にする働きをもつとし

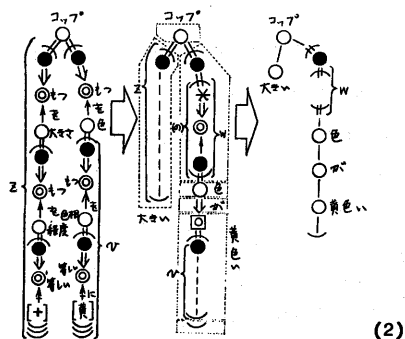
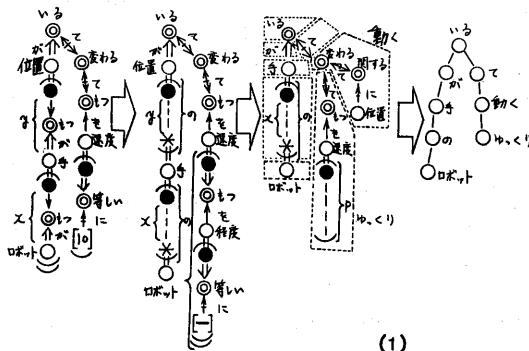


図8 単語分割の例

ているが、これは標準意味表現中には記されていない。

この為形容詞、形容動詞終止形の単語意味表現と比較する際には、その中の関係節又は名詞節中の動詞句と現在処理している意味表現とを比較する必要がある。その結果両者が包含関係にあれば、処理中の意味表現の該当部分を単語の意味に合せて関係節化又は名詞節化し、更に回又は回-を付加して形容詞化、形容動詞化する。

尚、処理中の意味表現をどのような細かさで単語分割するかは、単語記憶中にどのような語を用意しておくかによって決定される。以上の処理により意味表現が単語分割されると、そこから単語間の意味接続関係が決定される(図8)。

4.5 文生成

この段階までに作られた単語間の意味接続を表す木構造(図8)から表層文が生成される。この為に、表層の語順と単語の意味接続との間の関係を表す文型記憶(一例を図9に示す)が用意されている。処理手順の概略を以下に示す。

- i) 単語間の意味接続を表す木構造のheadの語の品詞を調べ、その品詞が支配する文型記憶を検索する。
- ii) 木構造のheadに係る各枝の先頭の語の品詞を調べi)で検索された文型記憶中のどの品詞、句、節に対応するかを調べて、表層上での順序を決定する。
- iii) 木構造のheadに係る枝が複数の語から成る場合は、その枝の各々について再帰的にi)からの処理を行う。

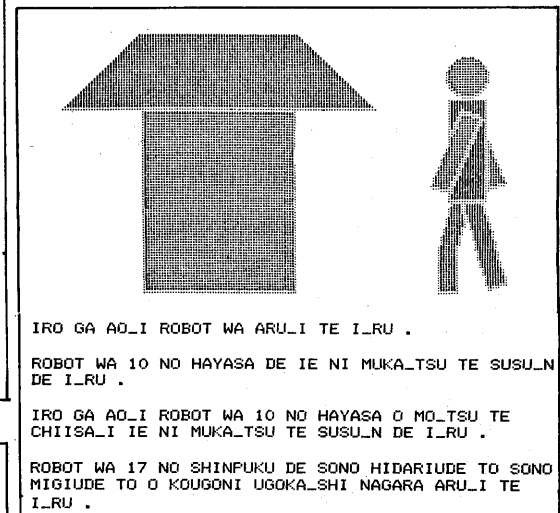
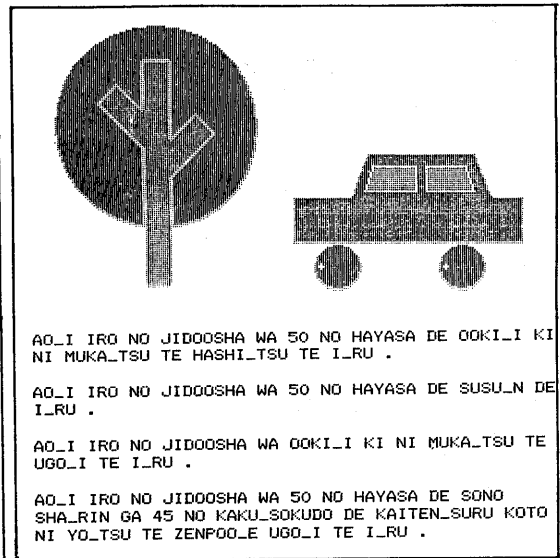
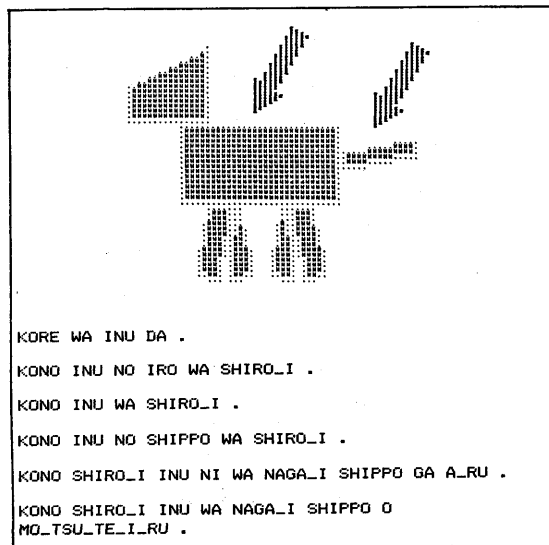
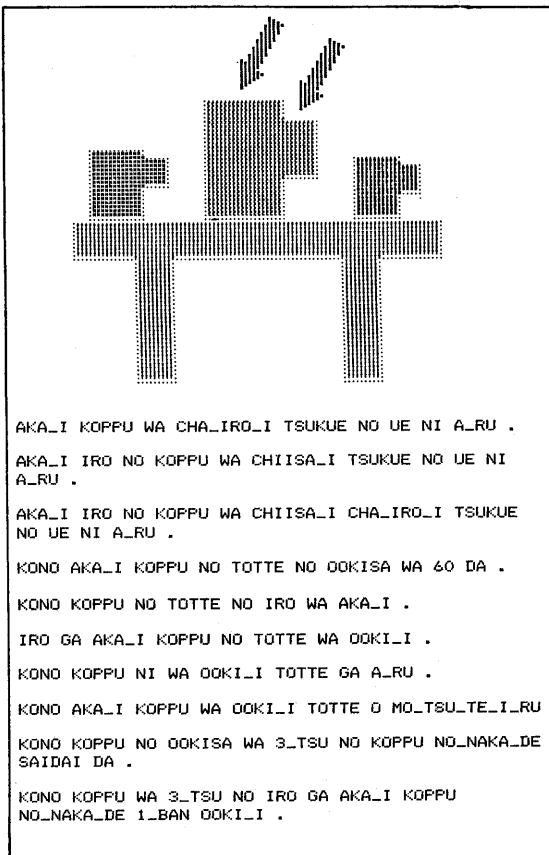
文出力の形式はローマ字の分ち書きとし、語幹と活用語尾等をハイフンで区切っている。尚、現在インプリメントされている単語数は232語である。



図9 文型記憶の構造

5. 動作例

以下に入力画面とそこから生成された日本語文の一例を示す。



6. まとめ

視覚情報から日本語文を生成するシステムについて報告した。本システムで採用した意味表現は表2の範囲内の英語文も同様に生成する能力をもつ。従って文型記憶を変えただけで、英語文を出力することが可能である。更に広範な文を生成出来る様に拡張することが今後の課題である。

(参考文献)

- 1) D.L.Waltz (1981) : Toward a detailed model of processing for language describing the physical world. IJCAI 1981
- 2) 曾我 他 (1980) : 言語・画像相互参照によるプロットの把握 信学技報 AL 80-85
- 3) 高木, 六沢 他 (1983) : 自然言語処理との共働を考慮した図形認識について 情報処学会 知識工学と人工知能研究会 (1983 2月)