手書き原稿認識における語彙および構文の検定†

池田克夫†† 大田友一†† 上野恵美子††*

分かち書きされていない日本文手書き原稿の認識において，文字認識の後処理として，単語および単語間の接続の二つのレベルで検定を行い，文字認識率を向上させることを試みた．文字認識の結果は，各文字位置に対する10個の候補文字として与えられる．この候補文字列中から，語彙辞書を用いて可能性のある単語を抽出する．さらに，単語間の接続条件を規定した接続テーブルを用いて接続検定を行うことにより，文法的に合法な文字列を構成しうる文字のみを拾い上げる．語彙辞書や接続規則の構成は，コード入力された文章の構文解析の場合と異なり，欠落や曖昧な文字を多数含む候補文字列との照合を必要とする，手書き文字認識処理の性質を考慮したものとなっている．本処理により，正解文字が第1位に現れる率である第1位認識率が，文字認識段階での第2～3位認識率にまで向上する．また各文字について10個ずつ与えていた候補文字を，1～3個程度に減らすことができるので，文字選択の操作が容易になると考えられる．

1. はじめに

今日，カナ漢字変換方式による日本語ワードプロセッサの普及がめざましいが，手書き原稿も依然多数書かれており，この文書処理の機械化も強く望まれている．文書の作成には推敲がつきものであるから，OCRとしては，文字認識率100%をめざして大変な努力をするのはあまり得策とはいえない．すなわち，日本語エディタによって文書処理を行うことが前提の場合には，ある程度の認識率が得られれば実用上差支えないと考えられる．この場合，文字認識段階では結果を1文字に限らずに適当な個数の候補文字を選び，その中から正しい文字列を選択することが実際的である．しかし，正解を含む10個以上もの候補文字をそのままオペレータに提示するのは現実的でなく，前後の文脈から正解でないことが容易に判断できるものはあらかじめ削除しておくことが適当であろう．

パターン認識において文脈を利用しようという考え方は，割合古くから示されている[1]．文字認識において文脈を利用するには，単語認識および単語間の接続検定という二つのレベルが考えられる．

単語認識は，文字列の確からしさを，単語辞書を用いたり文字の並びの文法的制約あるいは文字間の遷移確率を用いて検定しようとするものである．

認識すべき語の範囲が限定されていて，しかも孤立単語としてだけ認識すればよいようなメニューの入力などでは，単に単語辞書との照合を行うことにより，おおむね問題が解決する．阿部らは，単語が“分かち書き”されている英単語の場合に単語辞書との照合の効果について報告している[7]．このほかにも片仮名書き住所の認識など辞書を用いる研究の報告がなされている[8)~15)]．

長田らは，片仮名書きされた文字列を対象として，文字列が文節の並びであるという拘束を課し，その条件下での事後確率を最大にすることによって文字認識率の向上をはかる方法について述べている[16]．彼らは文字の図形的特徴と文脈との両方を生かして認識を行うために，両者を統一した認識尺度として文字列の生起確率を用いているが，文字列の生起確率を合理的に決定することは困難であるので一定値を用いている．

単語間の接続のレベルまでを文字認識に利用した例は上に比べて非常に少ないようである．新谷らは言語処理として単語処理と文節処理とを合せて利用する後処理方式を論じているが[17]，文節処理の文法についてはより詳細な検討を要すると思われる．

我々は，単語および単語列の二つのレベルにおいて検定を行い，文字認識率を向上させることを試みた．すなわち，**図1**に示すように，分かち書きされていない日本文の手書き原稿を対象として，文字認識の結果である候補文字列から，語彙辞書を用いて可能性のある単語（構文要素）を抽出し，さらに，これらの構文要素間の係り受けの検定を行う．我々の方法は，文章の意味論的な解釈には立入らず，文法的に合法な文章を候補文字列より抽出するものである．これによって，第1位認識率が文字認識段階での2～3位の認識

† Vocabular and Contextual Postprocessing for the Recognition of Handprinted Japanese Manuscript by KATSUO IKEDA, YUICHI OHTA (Institute of Information Sciences and Electronics, University of Tsukuba) and EMIKO UENO (Institute of Information Sciences and Electronics, University of Tsukuba. Currently with Nippon IBM Co.).

†† 筑波大学電子・情報工学系
* 現在 日本IBM（株）

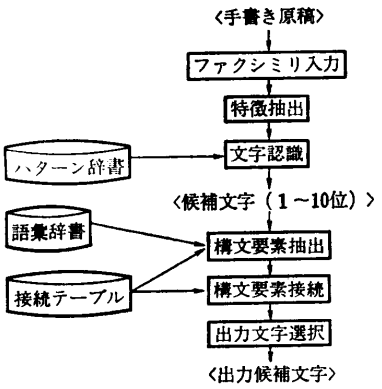

図 1 手書き原稿認識処理の流れ
Fig. 1 Flow of handprinted manuscript recognition.

率程度にまで向上することがわかった．また，各文字について10文字ずつ与えられていた候補文字を 1～3文字程度に減少させることができる．この結果，文字選択の操作が容易になるので，実用上有用である[2]．

本論文では，以下，第2章に日本文構文検定のための文法，第3章に語彙および構文の検定法，第4章に実験結果と考察を述べる．

2. 日本文構文検定のための文法

本論文では，日本文を構成する要素を構文要素と呼ぶことにする．構文要素とは，文を形成する最小単位である形態素が1個以上連結したものである．単語列のレベルでの検定は，文字認識結果の文字列中の二つの連続する構文要素間の係り受け関係の検定として行う．ここで，二つの要素間の接続関係は，先行する要素は後ろの要素に“係る”，後ろの要素は先行する要素を“受ける”と表現する．長尾らは同様の考え方により構文要素の接続を規定する文法を構成しているが[3]，これは，コードで入力された日本文の解析を目的としたものである．したがって，解析される文字列は“正しい”日本文であることが前提となっており，比較的“ゆるい”文法でも十分である．これに対し，我々の場合には，文字認識結果の文字候補中から文法的に正しい文字列を拾い出そうとするものであり，長尾らの場合よりはるかに“きつい”文法が必要となる．

日本語の文を構成する要素を適当なカテゴリに分けると，接続関係を規則の形で記述することができる．

- 大きな 赤い くつ
- 小さな くつ

という例を考えてみると，同じカテゴリに属すると思われる「大きな」，「小さな」の後ろには，「赤い」のようにさらに修飾が加わることもあるし，直接「くつ」に接続することもある．しかし逆に「くつ」を考えた場合には，その前に接続する要素は，「赤い」，「大きな」，「小さな」といった修飾語，あるいは，「の」，「が」などの助詞のように，かなり限定することができる．このように，日本文では，後の要素からの限定の方が，前の要素からの限定よりも強いことが多い．したがって，文法をできるだけ簡潔にし，正しい文章の探索を容易にするためには，後ろの要素から前の要素を規定する方向に働く文法を構成するのが得策と考えられる．

長尾らの構成した接続規則は，構文要素間の係り受けの条件として，ある要素に先行できる要素を，品詞と未然形・連用形などの活用の種類の組合せで分類して，接続カテゴリとしてまとめたものである．

我々の検討の結果では，接続条件を品詞と活用の種類だけで規定するのは不十分であることがわかった．たとえば，動詞「ある」は，形容動詞・断定助動詞・伝聞助動詞・様態助動詞の連用形，副詞，感動詞，接続詞，助詞，句点，読点を受けるとされている．しかし，詳細に調べると，ダ活用型形容動詞および様態助動詞の場合には，「ある」に係る連用形は受けるが他の連用形は受けない．また，格助詞「が・に」や，「か・しか・なり」以外の副助詞，および接続助詞「つつ」は受けるが，「…の（格助詞）ある」，「…ても（接続助詞）ある」などとはならない．

これらの例から明らかなように

1）活用の種類の指定には従来用いられている未然・連用・終止・連体・仮定・命令という6種の活用形では不十分である．

2）助詞は用法が複雑であり，個々の助詞を別個のカテゴリとして取扱う必要がある．

さらに，

3）五段活用動詞の語尾については，おのおの活用の行を指定する必要がある．

4）接続対象の指定には，集合を陽に指す方法と，補集合を指す方法とがある．後者は接続禁止を規定することに相当するが，集合の要素数が小さい方を指定すれば簡潔な表現が可能になる．

5）言語は，文法だけでは完全に記述しつくせないので，用例のごく少ない規則はむしろ慣用句として登録するのがよい．

我々は，以上の考察を基にして，より精密かつ簡潔な表現の文法を構成した．以下では，この文法の基本的な考え方を述べる．詳細については誌面の都合で割愛せざるをえなかったが，興味のある読者は文献 2) を参照されたい．

2.1 構文要素の属性

◎品詞：名詞，動詞，助詞などの品詞を表すために，工業技術院等で作成された品詞コード表[3),4)]を基本として，連続記号〔・〕，開始記号〔「etc.〕，閉鎖記号〔」etc.〕の品詞を加えた．

◎活用型：五段活用，上一段活用などの活用型を表すコードは，長尾の活用型コード表[3)]に準じたものを用いている．

◎活用行：五段活用の動詞の語尾が変化する行は，**表 1** に示したような 2 桁のコードを用いて表す．

◎活用段：活用段は，活用語の活用形を示すもので，13 種の活用段を指定する．**表 2** に示したように，未然形が四つの段に分かれているのは，

1）サ行変格活用動詞では「せる・れる」，「ない」，「ぬ」に係る活用形がすべて異なる，

2）五段活用動詞では，「ない」と「う・よう」に係る活用形が異なる，

など，未然形という一つの活用形では，4 種の語に係る活用形を規定しきれないためである．

また，連用形が三つの段に分かれているのは，

1）形容動詞では，「ある」に係るまたは中止の活用形，用言に係る活用形，「て・た」に係る活用形がすべて異なる．

2）五段活用の動詞と形容詞は，中止の活用形と用言に係る活用形は同じだが，「て・た」に係る活用形はそれと異なる．

などの理由による[6)]．

用言の語幹と，活用しない語にもそれぞれコードを与えた結果，計 13 種の活用段となり，ビットパターンにより各活用形が取りうる活用段を表現した．未然形・連用形に上記のような区別がない活用語の場合には，一つの活用形に二つ以上の活用段が対応することになる．たとえば，カ行五段活用の動詞「行く」は図 2 のように表される．

2.2 構文要素間の関係

構文要素間の接続関係を規定する規則を構成するため，構文要素をその品詞に応じていくつかのカテゴリ

表 1 五段活用動詞の活用行コード
Table 1 Codes for the categories of verbs with five conjugations.

コード	1	2	3	4	5	6	7	8	9
上位・音	清		濁		半濁				
下位・行	あ・わ	か	さ	た	な	は	ま	や	ら

表 2 活用段のコード
Table 2 Codes of thirteen conjugations.

活用形	コード	意味
未然形	1	「せる・させる」に係る
	2	「ぬ・ん」に係る
	3	「ない」に係る
	4	「う・よう」に係る
連用形	5	「ある」に係る，または中止用言に係る
	6	
	7	「て・た」に係る
終止形	8	文の終止
連体形	9	体言に係る
仮定形	A	「ば」に係って仮定の意を表す
命令形	B	命令の意を表す
	E	語幹
	F	活用なし

文字数	語	接続カテゴリ番号	品詞	活用型	活用行	活用段 0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1	行	01	19	11	12															1	
1	か	04	19	11	12		1	1	1												
1	こ	04	19	11	12					1											
1	き	04	19	11	12						1	1									
1	っ	04	19	11	12								1								
1	く	04	19	11	12									1	1						
1	け	04	19	11	12											1	1				

図 2 「行く」の属性表現
Fig. 2 Property expression of a verb "iku" which means "go".

に分け，そのカテゴリに属する要素が共通して“受ける”構文要素を拾い出し，この結果によりカテゴリの統合・分割を繰り返して，あるカテゴリに属する構文要素はすべて同じ属性の構文要素を受けうるようにした．このカテゴリを接続カテゴリと呼び，構文要素間の関係を表す属性として各要素に付与した．

接続カテゴリを形成する際に見いだされたいくつかの問題点と，その解決策を以下に述べる．

（1） 助詞の接続に関しては各種の例外的な用法が存在する．たとえば，副助詞の「は」は，格助詞「から」，「で」，「と」，「として」，「に」，「の」，「へ」，「より」，「を」を受けうるが，「も」を受けることはない．また「も」は，接続助詞の中では「たり」だけを受ける．このような係り受け関係は，助詞どうしに限らず，助詞と他の品詞との間にも存在し，品詞という大きな枠だけでは規定しきれない．そこで，助詞は識別コードを定めて個別に識別することにした．この助詞識別コードと，接続禁止の規定を用いることにより，たとえば，“「の」以外のすべての格助詞を受ける”という接続規則も，“格助詞を受ける”，“「の」は受けない”という二つの項目で容易に表現できる．

（2） 当初，名詞はすべて一つのカテゴリとしていたが，「〜のやり方」や「書き方」といった，動詞の連用形と名詞の接続は，すべての名詞に共通に許される接続ではない．この係り受けを一般に許すと，たとえば，「これら」の認識結果として，下一段活用の他動詞「こねる」の連用形「こね」と名詞「ち(地)」が接続した「こねち」などという類似の文字列が文法上すべて認められることになるので，動詞の連用形と名詞の接続は原則として禁止する．そのかわり，「方」，「性」などは，接尾語的名詞として名詞とは別のカテゴリを与えて動詞連用形を受けることにした．また頻繁に使用されると思われる言い回しは，複合語として辞書に登録する．

（3） 名詞は多くの品詞の語を受けるので，1文字の名詞の接続はできるだけ制限する必要がある．したがって，1文字名詞は三文字以上続かないものとし，そのような語が存在するならば，それは辞書に一語として登録することとした．

（4） 規則では検定しにくい例外的用法，および，唯一の用例は，接続カテゴリが明確となる要素までを含めた文節を一つの慣用句として取扱うことにした．この慣用句の接続カテゴリは最前部の構文要素のものを，品詞・活用型・活用行・活用段は最後部の構文要素のものを引継ぐ．

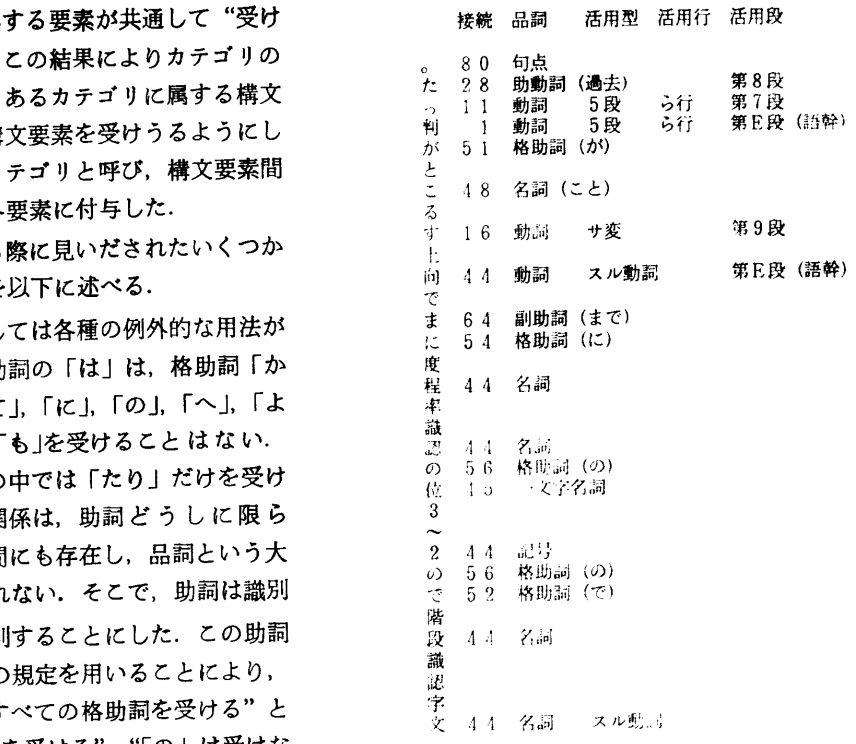

図 3 構文要素の接続例
Fig. 3 An example of morphological analysis.

接続規則は各接続カテゴリごとに接続可能項目と接続禁止項目を列挙した接続テーブルによって表す．各項目は，接続を可能または禁止とする構文要素の属性を規定するコードを連結したものである．我々の構成した文法では，接続カテゴリ数 86，接続可能項目数 353，接続禁止項目数 22 である．

図 3 に，この文法に従う構文要素の接続の例を示す．後述のように，文字列の後端より処理を行うため，文字列の順序を反転してある．

3. 語彙および構文の検定

3.1 語彙辞書

入力文字列中から構文要素を抽出するための辞書は，付属語テーブル・語彙テーブル・国語辞書の3種類を用い，この順に検索する．

◎ 付属語テーブル

助詞，助動詞，活用語の語尾，慣用句の計約 500 項目を格納している．検索を効率よく行うために木形式に格納してあり[5]，最長適合に至るまでに適合するすべての項目を抽出することができる．

◎ 個人用語彙テーブル

2,000項目程度までの個人用語彙を格納し検索効率を向上させようとするものである．このテーブルは動的に管理し，検定処理中も，国語辞書より抽出された語彙のうち文字認識における信頼度の高いものを順次格納していく．項目数が比較的少数に限定されているので特別の索引や構造は用意していない．

◎ 国語辞書

三省堂新明解国語辞典[6]を工業技術院でデータベース化したもの[3),4)]から10,400項目を抽出して使用している．ただし副見出しも一つの独立した項目となるように加工した．また，活用語は語幹のみを収容している．

後述のように，欠落を許した照合を可能とするため，順引き，および逆引きの索引を用意してある．漢字見出しについては先頭または末尾1文字，平仮名・片仮名の見出しについては，先頭または末尾から4文字までで索引を構成している．

3.2 候補文字列と評価値

文字認識の結果は，各文字に対して，第1位から第10位までの候補文字として与えられる．各候補文字には，入力手書き文字パターンと辞書パターンとの間の距離が付加されている．各文字に付加された距離の値は，候補順位の増加関数であり，文字の確からしさの評価値として使用するには不都合である．候補順位の減少関数としては，各候補文字が正解である確率を評価値に用いる方法が考えられる．このため，各候補文字の距離を候補順位第1位の文字の距離の値で正規化した値と，その文字が正解である確率との関係を実験的に求めて，距離から評価値への変換関数として利用している．

漢字の「一」や記号の「－」(マイナス，ハイフン，長音記号) などは，別個の文字として識別されなければならないが，形状だけからこれらを区別することはほとんど不可能なことが多い．これら，文字認識段階で明確に区別することが困難な文字の組に属する文字が候補文字に含まれている場合には，候補に含まれていない類似文字を強制的に候補に加える (最大15文字以内)．追加された類似文字，および元の候補文字に含まれていた類似文字で順位の低い文字の評価値は，上位の類似文字の評価値に基づいたプレミアムを加える．

3.3 構文要素の抽出

構文要素の抽出および係り受けの検定は文の後方から実行するが，全原稿の終端から処理するのは，処理結果の反転などの煩わしさをともなう．したがって，文章中の区切り点までを単位とし，その範囲で後から前へと処理する．もちろん，区切り点そのものも100％正しく認識されているとは限らないが，区切り点であることは高い確度で検出できるので問題はない．

区切り点で囲まれた20～40文字の文字列を入力とし後端より構文要素の抽出を行う．構文要素は，入力文字列中のすべての文字位置で終了するすべての長さの部分文字列を辞書項目と照合することにより抽出する．構文要素の候補がむやみに増大するのを避けるため，その要素の後ろに続く構文要素が受けうるもののみを抽出する．ただし，途中での誤りが伝播するのを避けるため，接続条件が満足されなくても，文字認識結果での候補順位が上位の文字のみから構成される構文要素は強制的に残すようにしている．

付属語テーブル・個人用語彙テーブル・国語辞書の順に検索し，いずれでも一致のとれなかった文字列については，英数字によって記号が構成されているかどうか検定を行う．

国語辞書は項目が非常に多く，見出し長が1文字の項目も多数存在するので，まともに照合を行うと，極端な場合はすべての平仮名が1文字ずつ独立の構文要素として抽出されるなどの問題が生じる．したがって，実際には，入力文字列の字種と長さに応じて国語辞書への参照を制限し，かつ，最長適合項目のみを抽出することにして，いたずらに多くの項目が取出されないようにしている．

部分文字列と辞書項目との適合は，一定割合の文字の欠落を許すことにする．最前端文字および最後尾文字の一方が欠落した場合にも対処できるよう，語彙辞書との照合は前方一致および後方一致をすべて試みる．このアルゴリズムでは，両方が欠落した場合には項目が抽出できないが，比較的認識率の高い場合には，このような確率はきわめて小さいとみなし，これを含めないことによる見逃しは無視する．

構文要素の候補として抽出された一続きの文字列には，構文要素としての評価値 E を与える．これは，文字列中の各文字の評価値の総和を基礎に，次の条件を満たすように変形したものである．

1) a，b を文字列とする．“・”を連結とすると，

$$E(a)+E(b)\leqq E(a\cdot b),$$

すなわち，短い構文要素二つよりも，長い構文要素一つの方に高い評価値を与える．

2) 欠落部分の文字の評価値は一定の正値とする．

3）評価値の高い文字列に，一定割合以下の評価値のきわめて低い文字が含まれているときには，その評価値 E を割増しする．

3.4 構文要素の接続と文字候補の出力

語彙辞書から抽出した構文要素を結合して，“文”とする．この場合も構文要素と同様，生成された文について評価値を求め，この評価値を基に，最終的に上位100位までの文を選択する．

接続が可能な構文要素どうしを結合して一続きの“文”を構成する場合に，生成される文の個数は，入力文字列の長さに対して指数関数的に増大する．したがって，文の後方から，ある程度接続検定が済んだ部分は，接続条件だけを残してどんどん出力することにして，処理系の記憶容量の指数関数的な増大を免れている．

構成した文を評価値の大きい順に取出し，各文字位置について最大3文字まで出力スタックに入れていく．このとき，各文字のもつ評価値を文字位置ごとに積算しておき，一定値を超えたならばその文字位置にはそれ以上文字を出力しない．

適当な語彙が発見されずに空白のまま，あるいは，積算した評価値が小さいままとなっている文字位置に対しては，積算値の大きさによって出力文字数を3，5，10文字の3段階に分けて，入力候補文字の中の評価値の大きいものから出力スタックに追加する．

4. 実験結果と考察

図4に本論文の方法によって処理した例を示す．上段が入力の候補文字列であり，下段が出力文字列である．出力文字列には，語彙の検定で正しい文字が選ばれた例と接続検定によって正しい構文要素が選ばれた例とを示している．ただし，入力文字列は分かち書きされていないので，語彙検定の際にも接続検定が関与し，この両者を明確に区別することはできない．当然のことながら印をつけたところ以外もすべて接続検定が行われている．

語彙および構文検定による文字認識率の改善効果を調べるため，約2,000文字の文書により認識実験を行った．パターン辞書としては，JIS 第1水準漢字を含む約3,200字種を用意した．文書原稿，パターン辞書ともに A4判400字詰原稿用紙に水性ボールペン（細字）で記入したものを16本/mmでファクシミリにより2値入力した．1文字は140×140 画素で量子化されている．文字パターンから抽出する特徴量としては方向寄与度密度特徴[18]と呼ばれるものを採用した．

語彙および構文検定により得られる認識率改善は，文字認識段階での認識率（入力認識率と呼ぶ）に依存すると考えられる．種々の認識率の候補文字列を用意するため，5人の筆者による同一内容の文書原稿と3種のパターン辞書により得られる15通りの組合せにより認識実験を行った．パターン辞書のうちの1種は文書原稿の筆者の1人と同一人によるものである．他の2種は，文書原稿の筆者以外の筆者による辞書パターンを混入して作成した．当然ながら，同一筆者間の組合せが最高の入力認識率を与える．

上記により作成された15組の候補文字列の認識率（入力認識率）と，それを本論文の方法で処理した後の認識率（出力認識率）の関係を**図5**に示す．図では，横軸が入力認識率，縦軸が出力認識率を示し，個個のカーブが一組の候補文字列による実験結果に対応する．各カーブ上の4個の印は，左から順に，第1位，第2位，第3位，第10位の認識率を示している．

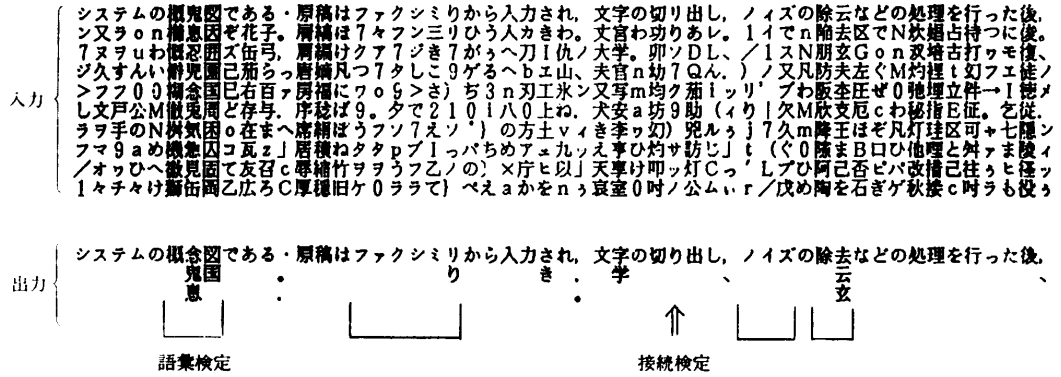

図4 システムの処理結果の例
Fig. 4 An example of the input and output strings.

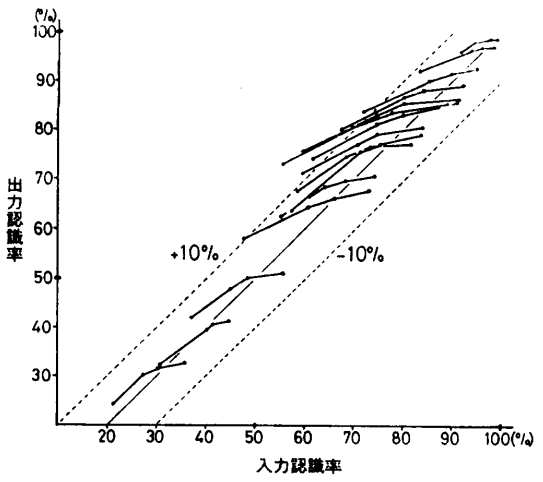

図 5 種々の入力認識率における認識率改善の効果
各カーブ上の点は一つの実験での1, 2, 3, 10位認識率を示す.

Fig. 5 Relation between the input and the output recognition rate.
Each curve shows the 1st, 2nd, 3rd and 10th recognition rate in one experiment.

斜め45度に引いた実線が改善率0％に相当し，それよりも上方にあるほど，改善率が高いことになる．この結果から次のようなことがわかる．

1）入力第1位認識率が50～80％で，第1位認識率と第10位認識率の差が20～30％ある場合には，最大の改善率が得られ，第1位認識率において10～15％改善される．

2）入力認識率が高い場合（第1位認識率90％以上），これ以上の改善の余地は少なく，第1位認識率において5％程度改善されるにとどまる．

3）入力認識率が低い場合（第1位30％以下），いくら文字の欠落を許しても，適切な単語が選択できず，第1位認識率において3％程度しか改善しない．

4）出力第1位認識率は，おおむね，入力の第2～3位認識率程度に向上する．

5）出力第10位認識率は，入力第10位の認識率よりも2～5％低くなる．これは，出力文字数を多くの場合1～3個に制限しているためである．すべての正しい解を構文的処理のみで拾い上げることは困難であることを示している．

5. 結　　論

手書き文字認識の結果得られる文字候補の中から，語彙および構文の検定を行い，より可能性の高い文字を選択することによって文字認識率を向上させる試みについて述べた．本方法は，日本文の意味論的な処理に立入らないために自ずと限界はあるが，より単純明解な論理で処理できることが特徴である．本論文では，二つの構文要素の係り受けのみを扱ったが，間に他の構文要素を挟み込むような，たとえば「必ずしも…でない」などの，慣用句を取扱えるようにすると，より有効となると考えられる．また，「文字認識」のような熟語も“「文字」を「認識」する”というように，主部と述部に分割して用いられることも多い．このような，三つ以上の構文要素の組合せに関する問題は今後の検討課題である．

謝辞　本研究は，文部省科学研究費特定研究「言語の標準化」および試験研究の補助を受けて行ったものである．本研究の実験に協力された本学技官鈴木秀則氏ならびに研究室各位に謝意を表する．

参 考 文 献

1) Toussaint, G. T.: The Use of Context in Pattern Recognition, *Pattern Recognition*, Vol. 10, No. 3, pp. 189-204 (1978).
2) 池田：日本文手書き原稿の浄書・編集システム，昭和58年文部省科研費報告書（1984）.
3) 長尾：計算機による日本語文章の解析に関する研究，文部省科研費報告書（1979）.
4) 荻野：国語辞書ファイル化作業，計量計画研究報告（1981）.
5) 長尾，辻井，山上，建部：国語辞書の記憶と日本語文の自動文割，情報処理，Vol. 19, No. 6, pp. 514-521 (1978).
6) 金田一編：新明解国語辞典，三省堂，東京(1974).
7) 阿部，秦野，福村：辞書を利用する文字認識系の能力の評価，信学論，Vol. 52-C, No. 6, pp. 305-312 (1969),
8) 宮崎，星野，立木：単語認識による文字認識の改善について，情処学会14回大会，pp. 229-230 (1973).
9) 星野，三浦，飯田：OCR のための単語認識，情処学会20回大会，pp. 487-488 (1977).
10) Doster, W.: Contextual Postprocessing System for Cooperation with a Multiple-choice Character-recognition System, *IEEE Trans. Comput.*, Vol. C-26, No. 11, pp. 1090-1101 (1977).
11) 辻，江上，森：英文住所の認識，情処学会23回大会，pp. 683-684 (1981).
12) 印牧，中島，荒川：単語辞書を活用した文字認識法の一検討，信学技報，PRL 81-91, pp. 69-76 (1981).
13) 飯田，杉村：パターン認識における単語照合処理の一検討，信学技報，PRL 82-77, pp. 93-98 (1982).

14) 蕪山他：手書き漢字認識における単語辞書の利用，信学総合全大，p. 1341 (1982).
15) 土井他：文脈を活用した手書き語の認識，信学技報，PRL 73-73, pp. 1-10 (1973).
16) 長田，牧野，日高：日本語の文脈情報を用いた文字認識，信学論，Vol. J 67-D, No. 4, pp. 520-527 (1984).
17) 新谷，梅田：複合後処理法による文字認識精度向上の評価，信学技報，PRL 83-42, pp. 25-34 (1983).
18) 萩田，内藤，増田：大局的・局所的方向寄与度密度特徴による手書き漢字認識方式，信学論，Vol. J 66-D, No. 6, pp. 722-729 (1983).

(昭和59年11月12日受付)
(昭和60年 2 月21日採録)