

音声入力における誤認識訂正処理

三橋和男, 八田敏, 平塚良治 (沖電気工業株式会社)

1. まえがき

近年, 計算機が対象とする業務が拡大し, それに伴って計算機が扱うデータも, 従来の数値データに加えて, 日本語文字データ, イメージデータ等へと拡大してきている。

日本語の文書を扱う処理では, 欧米の26文字と違って数万種類もの文字を対象としなければならず, その入力方法を容易にすることが, 重要な課題である。

現状の日本語入力装置では, 入力方式として, 読みがなをキーボードから入力するかな漢字変換方式, タブレットによる入力方式, 2ストローク入力方式などマニュアルによる方式が中心であるが, 音声やOCRなどの認識装置を利用した入力は操作方法が容易なこともあり, 将来の有望な日本語入力方法と考えられる。

しかし, 現状の音声認識技術では, 次の欠点がある⁽¹⁾とされている。

(1) 日本語文書の入力の場合, 単音節単位の入力となり発声が不自然で入力速度が速くならない。

(2) 音声認識装置の認識率は, 必ずしも100%ではなく, 認識音節の中に若干の誤り音節がある。

以上の2点の問題を解決することが, 音声入力方式の普及につながると考える。

(1)については, 単音節入力での認識率が90%の場合, 約40字/分程度の入力が可能であると報告されており⁽²⁾, 日本語処理を施すことより, ある程度自然な入力が可能と思われる。(2)については, 現在いくつかの有効な方法が提案されているが⁽³⁾, 決定的な方式が確立している訳ではない。本稿では, (2)の誤認識音節を訂正する方法の一つとして, 入力した文節中の付属語部分の訂正方式について述べる。

2. 音声入力用かな漢字変換

音声による日本語入力方式では, 現在日本語入力手段として最も一般的であるキーボードから読みがなを入力し, 漢字かな混り文に変換する方式と基本的に同じ方式を採用している。しかし, 音

声入力用として独自の処理も必要である。本章では, 音声入力用かな漢字変換について述べる。

音声入力におけるかな漢字変換とキー入力におけるかな漢字変換の相違を簡単に示せば, キー入力の場合は読みが1個なのに, 音声入力の場合には, 読みがなが複数個あることである。しかし, その複数個の読みがに何らかの順位を付けることができれば, それ以降のプロセスはキー入力の場合と等しくなる。そこで文節内に n 個の文字があれば, そのときの候補文字列は 2^n 個(1音節発生に2個の候補音節が渡されるとき)になる。この 2^n に対して, 日本語の文字列のつながりを考慮して順位をつける処理が音声入力用かな漢字変換の特徴と言える。特徴を以下に示す。

(1) 最長一致法による読みがな抽出

認識された候補音節行列の中から, 文節となる音節列を抽出する処理において, 単に認識率の値だけでなく, 文節かな漢字変換で使用されている最長一致法を採用することにより, 正確な読みがな抽出を行う。

(2) 見出し辞書による処理の高速化

かな漢字変換において, 単語辞書の果たす役割は大きい。読がな列の中から, 文節を構成する自立語を引き当てたり, その自立語の品詞, 活用形等を単語辞書より得て, 後続する付属語との文法チェックを行う。

一般に単語辞書は, 1単語単位に

- ・見出し語
- ・漢字表記
- ・文法情報(品詞・活用形・等)

のデータを持ち, 一般には, 単語の配列は見出しの「アイウエオ」順に格納されている。

音声入力かな漢字変換でも, 入力された認識音節候補の行列から文節を構成する読みがなを抽出するときに, 単語辞書で処理が可能である。しかし, 補数が 2^n あるため, 辞書アクセス回数が増え, 非常に多くの時間がかかる。我々は, 音声入力に適した見出し辞書を考案した。

見出し辞書は単語辞書に格納されている単語の見出しを集めたもので, 次のような配列の方法を採用し, 必要最小限の辞書アクセス回数で済むよ

うに工夫している。

- ・見出し長単位に配列する
- ・見出し長が等しい時は母音コード順に配列する
- ・母音コードが等しい時はアイウエオ順に配列する

さらに、単語辞書をアクセスせずに見出し辞書だけで、文法的接続のチェックが高速にできるように見出し語に含まれる各単語の文法情報(品詞活用形・等)を見出し辞書に付加した。

3. 音声認識誤り訂正処理

音声認識誤りの訂正には2レベルある。第一のレベルは長音訂正とか、話し言葉から書き言葉への変換を主としたもので、その訂正の方法はすでに確立していると考えられる。第2のレベルでは認識された音節候補にない文字の訂正を含むもので、方法やアルゴリズム等、技術的にも未確立のレベルのものである。

これから述べる訂正処理は第2のレベルのもので、訂正のための前提条件として文節内での認識の誤りは高々1文字のみであるという仮定のもとで実行するアルゴリズムである。

我々のシステムでは発声の単位を日本語文の基本的単位である文節単位としており、その構成要素は、自立語と付属語列に分解できる。文字が正しく認識されたかどうかは自立語が辞書にあることと、自立語と付属語列が正しく接続できるかどうかで判断される。認識された文字に対しての処理の概略を図1に示すとともに、処理の内容を以下に示す。

(1) 文節構成が正しいかどうかのチェック

もし利用者の発声した音節が正しく入力されていれば、認識した文字列で正しい文節が構成できる。それを確かめるために、認識された文字列の先頭から最長一致法を利用して辞書を検索して、自立語を決定し、残りを付属語列とする。その両者が接続可能であれば正しい文節構成と考え、特に訂正は行わない。

(2) 付属語に誤りがあるかどうかのチェック

(1)で正しい文節が構成できない場合、最長一致法で見つけた自立語は正しいとし、付属語に誤りがあると考え、付属語列を訂正してみて正しい文節が構成できれば、その文字が誤りであるとする。

(3) 自立語に誤りがあるかどうかのチェック

(1)、(2)で正しい文節を構成できる文字列がない場合、自立語に誤りがあると考え、全体の文字列の後方から付属語を探し、マッチングするものがあれば、その文字列より前で自立語を構成するものとして自立語の文字訂正を試みる。その結果、正しい文節が構成できれば、その文字が誤りであると考えられる。

(4) 複数文字が誤っている時

(1)~(3)のいずれでも正しい文節が構成できない場合は、複数文字以上の誤りとして訂正しないものとする。単音節の認識率が95%程度であれば、1文節内で複数文字が誤る確率は数%程度以内と考えられるので、処理の対象とはしない。

(5) 自立語訂正と付属語訂正

今回の我々の提案する方式では自立語の訂正も付属語の訂正も全く同様のアルゴリズムで可能であるし、(2)の処理と(3)の処理の順序を入れ換えて、自立語を優先させることも可能である。しかし、次に示す理由によって付属語を優先させることとし、自立語の訂正は将来の検討項目とした。

(a) 個数面から

我々が所有している辞書の自立語数は約3万5千単語程度で、付属語接続語数約1300の30倍程度になり、付属語訂正ほど顕著な効果は期待できないと考えられる。

(b) 頻度面から

付属語接続語も一般の単語と同様に考えると、約1300単語が全文節の約55%以上⁽⁶⁾に出現すると解釈できるので、付属語は最も頻度の高い単語群と考えられる。

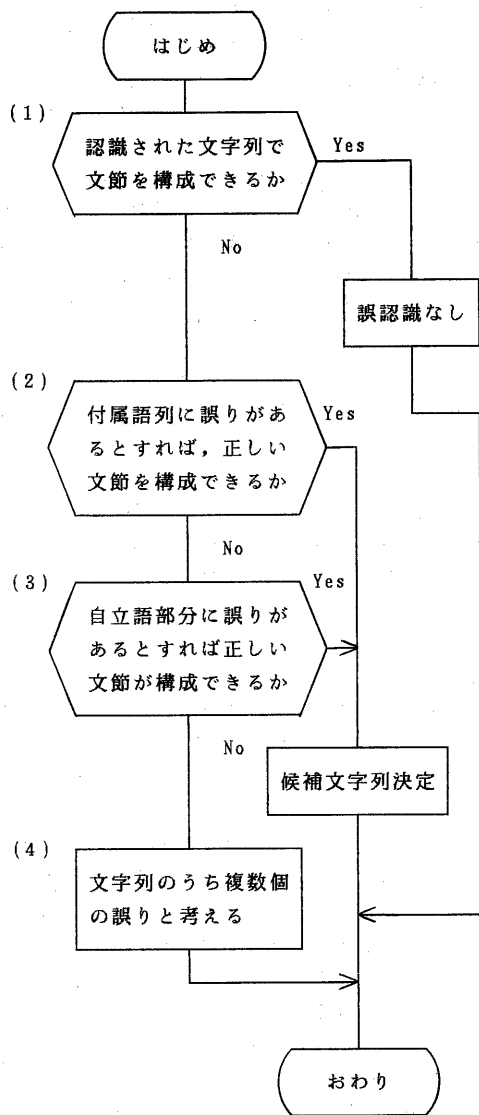


図1 音声認識誤り訂正処理

4. 付属語訂正処理

入力された文節の付属語文字列の誤り訂正は、まず、入力文字列で付属語辞書を検索し、辞書内に、入力文字列と同じ付属語が存在し、且つ自立語との文法的接続性が認められれば、入力文字列に、誤りがないこととする。その逆に、辞書に該当する付属語がないとき、または、付属語が存在しても文法的に接続性が認められないときは、入力文字列に誤りがあると判断し、付属語訂正処理を行う。

訂正処理は、入力文字列の代わりに、付属語辞書の中から、文節を構成している前半の自立語と最も合致する付属語を抽出することである。図2に付属語訂正の概略フローチャートを示すとともに、以下に付属語訂正処理の内容を説明する。

(1) 母音コードによる付属語抽出

単音節音声認識装置の認識特性として、音節を構成する(子音+母音)の内、母音の誤りは、ほとんどなく、誤りは子音にあることを利用する。付属語辞書内に入力文字列と同じ付属語がないときや、あっても文法的に接続性がないとき、入力文字列と同じ母音の付属語を抽出する。この中に正しい付属語が含まれている。例えば、入力文字列が「ナグテ」とすると、辞書から同一母音の語として、「タクテ、ダッテ、チャッテ、ナクテ、マスネ」の5語を抽出することができる。以上の処理を高速に行うため付属語辞書の配列を、従来の「アイウエオ」順でなく、音節長単位に母音の等しい単語を一箇所にまとめる必要がある。そのため、母音コードを設定し、付属語の読みを母音コード化し付属語列を母音コードの昇順に格納した。母音コードは、表1に示す様に「ア」段に対して1、「イ」段に対して2、「ウ」段に対して3、「エ」段に対して4、「オ」段に対して5、「ン」段に対して6(3)を対応させコード化したものである。付属語辞書の例を表2に示す。

(2) 1文字誤りの付属語抽出

母音コードの同じ付属語列が複数あるときは、その中から、入力文字列と異なる文字数が1文字以内の付属語を抽出する。これは、音声認識装置の認識率と、文節を構成する文字数 n (n の平均を5~6文字とする)を考慮すると、文節内に誤りがあっても、1文字程度と考えられる。単音節認識率95%の装置における文節の認識率を図3と表3に示す。文節の音節長が8音節として計算

しても1文字も誤らない確率は66%であり、1文字だけ誤る確率は28%である。従って文節中1文字以内の誤り確率は94%となる。よって、8音節程度の文節では1文字だけの訂正が適当と判断できる。

(3) 文法的接続性のチェック

処理(1)、(2)により抽出した付属語候補を、それが接続する自立語の品詞、活用形で文法的接続性をチェックし、接続性が認められた付属語を抽出する。抽出された付属語が複数個あっても、それらは以降の処理では優先順位付けはなされるが、訂正付属語としての候補からは削除されない。しかし、逆に(1)、(2)の処理で訂正付属語の候補として1個しかない付属語でも、文法的接続性が許されないものは、候補として却下される。文法的接続性のチェックの内容を図4のフローを参照しながら説明する。

(a) 自立語処理

自立語見出し辞書より、入力文字列の自立語部分の見出しを最長一致法で検索し、見出しの品詞、活用形を得る。用言のときは活用語尾処理を行う。見出し辞書も2章で説明したごとく、見出しの音節長単位に、母音コードの昇順に格納され、見出しに含まれる各単語の品詞、活用形等の文法情報がコンパクトにセットされている。

(b) 活用語尾処理

自立語が用言の場合は活用語尾辞書と照合し、活用語尾を付加する。

(c) 付属語処理

(a)で得た自立語の品詞・活用形と付属語が接続可能な自立語の品詞・活用形を照合し、合致するときは、文法的接続性を認める。

例えば、「マチマズ」と入力があった場合、最長一致法による見出し辞書検索により、「マチ」(名詞)が見つかる。次に付属語の「マズ」を辞書で照合するが該当する付属語がない。「マズ」と同じ母音コードで「マズ」と1文字違いの付属語は、「マス」(助動詞)がある。しかし、名詞と助動詞「マス」の文法的な接続性はないので日本語としては不適當である。このような場合は「マチ」の末尾からもう1字削って「マ」という自立語を調べる。自立語見出し「マ」は、ワ行五段活用(舞)、夕行五段活用(待)、名詞(間)

等の異なる品詞の見出しと解る。ここで、後の文字列「チマズ」、「チマス」との文法チェックをすると、「チ」は夕行五段活用の連用形の活用語尾、「マス」はそれに接続する助動詞と判断し、「マチマス」が、文法的に接続可能な自立語・付属語の接続となり、結果的には付属語「マズ」が「マス」に訂正され「待ちます」という文節に変換されることになる。

(4) 頻度と認識特性表による優先順位設定

文法的接続性チェックにより候補に残った付属語が1個であれば、それが正しい付属語として訂正処理が完了する。候補として複数個の付属語が残ったとき、その中から訂正付属語を1個だけ抽出することは、文全体の意味処理を行わなければならない。そこで、抽出された付属語の頻度情報と認識装置の認識特性を利用して、抽出した訂正付属語の候補に優先順位を設定し、訂正率の向上を図る。

Confusion Matrixは、101音節を発声し出力された認識結果を表にしたもので、表4にその例を示す。例えば、「ナ」と発声したときに、「マ」「ニャ」の順に誤り易いことが解る。入力文字列が「ヤマガラ」とすると(1)~(3)の処理で「山から」、「山なら」、「山やら」の3個の付属語が抽出される。この中で、「ガ」と誤認識されやすいのは、Confusion Matrixの出力軸の「g」を下方に検索すると「k, t, r」になる。つまり、「カ、タ、ラ」を発声すると「ガ」に誤ることがあることを示している。よってConfusion Matrixを利用すると、「山から」の「から」が最優先の訂正付属語となる。

また頻度情報から優先順位を決定すると、

1位	「から」	頻度	2410
2位	「なら」	頻度	399
3位	「やら」	頻度	27

となり、「山から」が最優先となる。このように両面から優先順位を決めることができるが、我々は、両者の積を計算し、その大小により優先順位をつけている。

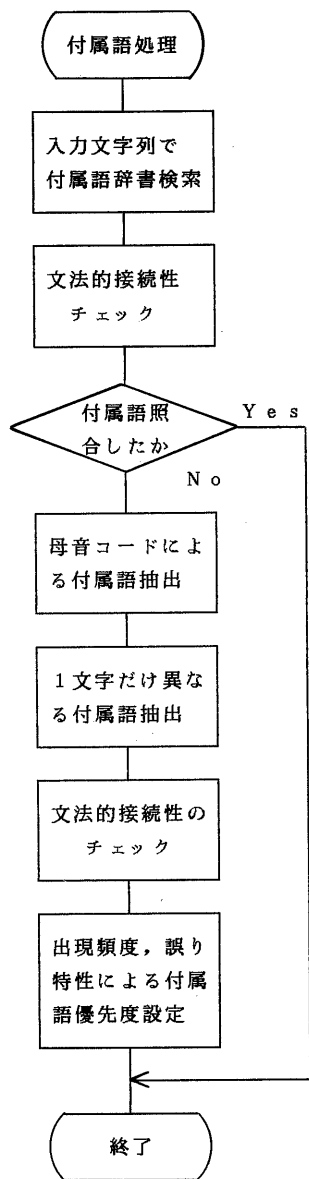


図2 付属語処理フロー

表1 母音コード表

母音	コード
a	1
i	2
u	3
e	4
o	5
n	6 (3)

表2 付属語辞書

読み	母音コード	度数	文法情報
タルニ	1 3 2	1	—
タクテ	1 3 4	1	—
ダッテ	1 3 4	1 1	—
チャッテ	1 3 4	1	—
ナクテ	1 3 4	4	—
マスネ	1 3 4	4	—
マスト	1 3 5	2 0	—
マスヨ	1 3 5	9	—

表3 単音節認識率 $i = 95\%$ における
文節長 n での誤り音節数 r の確率

r \ n	2	3	4	5
0	0.9025	0.8574	0.8145	0.7738
1	0.0950	0.1354	0.1715	0.2036

r \ n	6	7	8	9
0	0.7351	0.6983	0.6634	0.6302
1	0.2321	0.2573	0.2793	0.2985

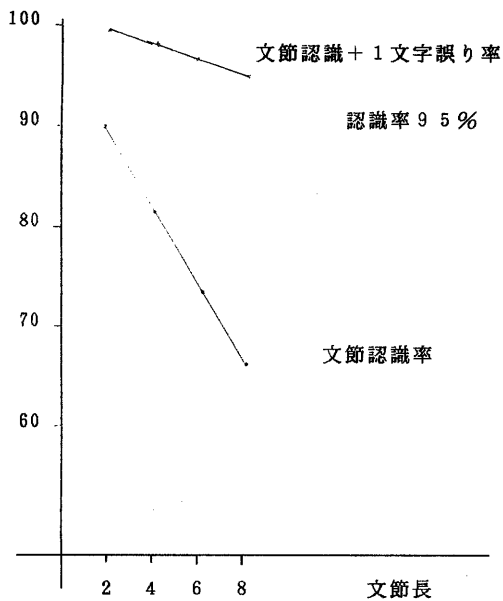


表4 コンフュージョンマトリックス

出力 \ 入力	V	k	s	t	n	h	m	y	r	w	N	g	z	d	b	p	ky	sy	ty	ny	hy	my	ry	gy	zy	by	py		
V	90																												
k		88		1								1																	
s			89									1																	
t				88							1	1																	
n					80		8														2								
h		2				88																							
m					3		87																						
y								54																					
r						1			85			1	1	1	1														
w										18																			
N											18																		
g												89		1															
z				1									89																
d														54															
b									2						82	6													
p					1											5	84												
ky																		53							1				
sy																			54										
ty			1																	49						3			
ny							1														50	2							
hy																	1	1			50				2				
my							1														3	49							
ry																							49	3		2			
gy																								49	3		2		
zy																									1	52			
by																										4			
py																											41	9	
																											2	5	47

図3 単音節認識率 $i = 95\%$ における
文節長 n での誤り音節数 r の確率

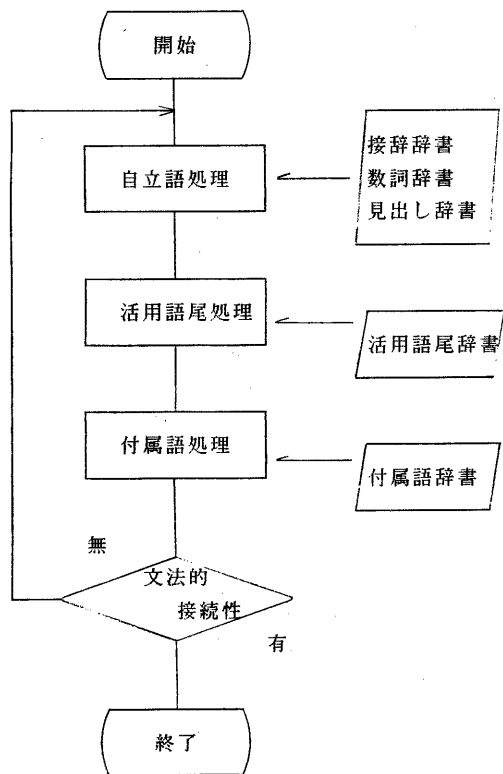


図4 文法処理フロー

5. 調査結果

我々が提案した付属語処理の効果を，電子計算機による新聞の語彙調査(3)(国立国語研究所)の助動詞・助詞連接表の付属語1297語をデータとし，調査した結果を示す。

(1) 母音コードに付属語抽出

母音コードによる付属語候補の絞り込みは，表5に示す様に，付属語の文字数が長い程絞り込みの効果が高く，7文字以上の付属語であれば，母音コードだけの処理で訂正が可能となる。

(2) 付属語訂正処理(1)～(3)を実施した場合

付属語訂正処理(1)～(3)による付属語抽出の候補数は，表5に示す通りになる。訂正付属語の候補は，全付属語数のうち，60%を一意に決定ができ，候補を3語まで許すと，95%まで達する。

(3) 頻度情報処理

表5で示す通り，付属語を訂正した場合，候補付属語数が4文字以上のものが21種類ある。しかも1文字付属語の中には4種あり，1文字付属語は文章中の出現率が高いため，問題があると思われる。しかし，頻度情報で優先順位を決定すると，上位3位までの訂正付属語が正しい確率は頻度数から換算して96%となり，問題点をカバーできる。表6に付属語文字数単位に，上位3個までの訂正付属語の出現率を示す。

表6. 候補3位以内付属語の出現率

付属語文字数	候補3位以内付属語の出現率
1文字付属語	96.15%
2文字付属語	99.69%
3文字付属語	99.22%
4文字付属語	99.71%
5文字付属語	99.72%
6～8	100%

表5 付属語訂正処理による抽出結果

付属語文字数	付属語候補数	母音コードによる分類	付属語訂正処理(1),(2),(3)による抽出
1文字 (31)	1	1	13
	2	0	1
	3	1	1
	4	1	1
	5	0	1
	6	2	0
	7	0	1
11	1	0	
2文字 (167)	1	4	56
	2	2	39
	3	5	15
	4	3	2
	5	1	2
	6	0	1
	7	3	0
	8	1	0
	9	1	0
	10	2	0
	12	3	0
13	1	0	
20	1	0	
3文字 (345)	1	30	170
	2	14	45
	3	8	17
	4	8	8
	5	8	1
	6	6	0
	7	4	0
	8	4	0
	9	3	0
	11	1	0
12	1	0	
13	2	0	
19	1	0	

付属語 文字数	付属語 候補数	母音コード による分類	付属語訂正処 理(1),(2),(3) による抽出
4 文 字 (368)	1	8 6	2 2 5
	2	3 5	6 6
	3	2 3	5
	4	1 1	3
	5	6	0
	6	6	0
	7	3	0
	1 2	1	0
5 文 字 (238)	1	1 4 8	1 9 2
	2	2 7	2 0
	3	8	1
	4	3	1
6 文 字 (100)	1	8 0	9 0
	2	7	5
	3	2	0
7文字 (33)	1	2 9	3 3
	2	2	0
8文字 (12)	1	1 2	1 2
9文字 (3)	1	3	3

自立語の場合は、単語数が非常に多く、今回調査した付属語の語数の30～50倍ある。付属語と同一の処理では、訂正した単語の候補が多すぎるので、ある規準を設け候補単語を絞り込む必要がある。

7. 参考文献

- (1) 中田和男：音声情報処理技術の動向，情報処理学会誌，Vol. 24, No. 8, p. 968～975, 1983
- (2) 並木，浜田，中津：音声認識を用いた日本語入力方式，通研実報，Vol. 32, No. 11, p. 2, 293～2, 303, 1983
- (3) 国立国語研：電子計算機による新聞の語彙調査(3) 1972 秀英出版
- (4) 栗田，相沢：日本語に適した単語の誤入力訂正法とその大語 単語音声認識への応用，Vol. 25, No. 5 情論節, p. 831～841
- (5) 高木，村田，小林：単音節音声認識装置の開発，沖電気研究開発122, Vol. 51, No. 1, 1984
- (6) 平塚，八田，津田：べた書きかな漢字変換の一方式について，日本文入力方式13-4, 1984

6. おわりに

(1) 結論

我々が提案する付属語訂正処理は、調査結果より、音声認識装置が誤認識した文節中の付属語の誤りの内、60%以上を一意に訂正が可能であり又、自立語の同音異義語と同様に、訂正した付属語にも候補を2語まで許せば85%、3語までで96～99%以上の訂正率を得ることができる。

(2) 今後の方針

誤認識した文字の訂正は、付属語についてのみ行っているが、今後は、自立語への展開を行う。