九州大学学術情報リポジトリ
Kyushu University Institutional Repository

日本語統計情報の音声認識への応用

高木, 英行
松下電器中央研究所

中嶋, 章子
松下電器中央研究所

楠原, 久代
松下電器中央研究所

前原, 文雄
松下電器中央研究所

https://hdl.handle.net/2324/4483175

出版情報：日本音響学会講演論文集, pp.41-42, 1985-03. The Acoustical Society of Japan
バージョン：
権利関係：

KYUSHU UNIVERSITY

1—4—21　　日本語統計情報の音声認識への応用*

○高木英行　△中嶋章子　楠原久代　前原文雄　（松下電器　中研）

1.まえがき　音声入力により文章を作成する、いわゆる音声WPにおいて、音声の誤認識の訂正や書き言葉・話し言葉変換を言語情報を利用して行うことが試みられている[1]。かな漢技術を利用した構文レベルの処理が中心となっているが、認識候補に直接構文レベルの処理を行うと言語処理量が増大する。そこで構文レベルの構文レベルの前処理として、種々の日本語の統計的特徴情報を利用して候補の予備選択を行い、構文レベル以降の言語処理量を低減することが有効となる（図1）。ここでは、日本語統計情報の特徴とその効果について報告する。

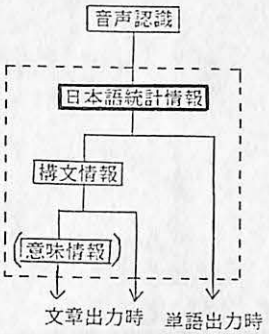


図1　言語処理を伴う音声認識系

2.日本語統計情報の特徴　日本語の特徴を以下3情報の観点から調べる。調査対象は約3万語の自立語辞書、約250語の付属語表から得られる全文節である。

2.1 音節連鎖情報[2][3]　音節間の接続には片寄りがあり、音節の全組合せの約半分しか接続せず認識候補削減に有効である。音節連鎖を調べる上で最も重要な事はかな漢字処理対象の全日本語を網羅しているかどうかである。従って多数文章からではなく、構文情報から考えられる全品詞組合せから情報抽出を行なわなくてはならない。音声WPのキー入力も考慮して（101音節＋かな文字）の拡張音節で調べると、音節連鎖の存在は全音節組合せの40.4%である。

2.2 同音異文字変換情報　音声で［zi］［zu］が入力されても文字の表記が一意に決定できない。促音入力ができない単音節認識の場合は［tu］も同様である。2通りの可能性を考慮して辞書検索すると処理速度に影響する。ところが「ぢ／づ／促音」は音節間接続、特に後接音節との接続に特徴があり、101単音節中、表1のように限られた音節にしか接続しない。例えば、「実に」を単音節入力した時、音節連鎖チェックをパスした［zituni］に対し「じつに」「じっに」「ぢつに」「ぢっに」が同音異文字候補となるのに対し、表1の情報で「じつに」以外は棄却される。自立語辞書・一般文章における効果を表3に示す。この情報により同音異文字音節の3～4割が一意に表記文字を決定できることがわかる。

表1　同音異文字変換情報

同音異字音節	後続可能音節数	例外
ぢ	13種	鼻血 添乳
づ	23種	固有名詞 （焼津など）
促音	35種	なし

2.3 長母音変換情報　「相当」が「そうとう」「そうとー」「そーとう」「そーとー」となる様に、長母音化読みを辞書に登録するとメモリ量の増大が甚だしい。逆に（エ段＋長母音）,（オ段＋長母音）の場合、各々（エ段＋「い／え」）,（オ段＋「う／お」）の可能性を考慮して辞書検索すると処理速度に影響する。ところ

* "Application of Japanese Linguistical Statistics to Speech Recognition," by H.Takagi, A.Nakajima, H.Kusuhara and F.Maehara(Matushita Electric Ind.Co.,LTD.)

が長母音と前接音節の間には特徴があり、長母音が「お」になり得るのは、前接音節が7種、「え」になり得るのは2種しか存在しない（表2）。しかも「ね+長母音」が「ね+い」になることはない。例えば、「壮麗」を音声入力した時、「そーれー」に対し「そうれい」「そうれえ」「そおれい」「そおれえ」が長母音変換候補となるのに対し、表2の情報で「そうれい」以外は棄却される。自立語辞書・一般文章における効果を表3に示す。（エ段+「い/え」），（オ段+「う/お」）が全て長母音化するわけではないが、この情報により長母音の約7割が一意に文字の表記を決定できると推測される。

表2　長母音変換情報

変換	前接可能音節
長母音 →「お」	お、こ、ご、と、ど、の、ほ
長母音 →「え」	ね、べ

表3　同音異文字変換・長母音変換で一意決定できない音節数

		名詞類	自立語	一般文章
全体数		62550	84138	38428
じ/ぢ、ず/づ、つ/っ	表1の場合	2472	3418	1168
	全ての場合	3569	4890	1883
エ段 +（い，え）	表2の場合	1429	1748	885
オ段 +（う，お）	全ての場合	4980	5698	3159

3.日本語統計情報の効果　図2の実験系において一般文章を音声入力したものとし、長母音の誤認識はないものとして単音節認識率70，80，90%の認識候補を第5位まで出力する。認識信頼度に基づいて単音節候補から文節候補を作成し、①～③の処理部で各日本語統計情報による候補削減を行う。その後正解との照合を行う。正解文節と照合されるまでの回数は、即ち構文レベル以上の処理回数を意味する。日本語統計情報を順次利用した時の効果量を図3に示す。

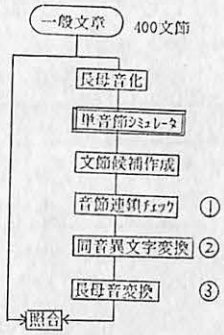
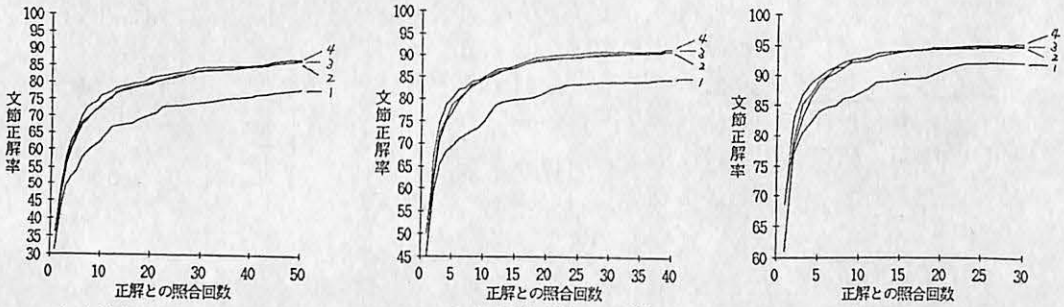


図2　評価実験構成



（a）単音節認識率　70%　（b）単音節認識率　80%　（c）単音節認識率　90%

1：日本語統計情報を利用しない，2：音節連鎖情報を利用，3：音節連鎖情報と同音異文字変換情報を利用，
4：音節連鎖情報と同音異文字変換情報と長母音変換情報を利用．

図3．一般文章における日本語統計情報の効果

4.まとめ　構文レベル以上の言語処理量低減のための前処理として音節連鎖情報・同音異文字変換情報・長母音変換情報を取上げ、その特徴を明らかにし文章作成時における文節訂正率と言語処理量との関係を求めた。これら3情報のうちでは音節連鎖情報が最も有効であり、次いで長母音変換情報が正解が文節候補の比較的上位にある場合に有効であることが明らかとなった。また同音異文字変換情報は目立った効果が見られなかったが、今後評価用文節データを多くして再評価すると共に、他の日本語統計情報について検討していく予定である。

参考文献
1) 河田他：情報処理学会第25回全国大会，6H-7
2) 外川他：音講論 2-7-4　昭57.10月
3) 坪井他：音講論 1-3-6　昭59.3月

謝辞　日頃から御討論いただいている中研各位に感謝します。