

## WEB 文書を利用した音声認識誤りの訂正方法の検討

西崎 博光<sup>†</sup> 伊藤 友裕<sup>††</sup> 関口 芳廣<sup>†</sup> 中川 聖一<sup>†††</sup>

<sup>†</sup> 山梨大学 大学院医学工学総合研究部

<sup>††</sup> 山梨大学工学部コンピュータ・メディア工学科

〒 400-8511 山梨県甲府市武田 4-3-11

<sup>†††</sup> 豊橋技術科学大学 情報工学系

〒 441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1-1

E-mail: <sup>†</sup>{[nisizaki,sekiguti](mailto:nisizaki,sekiguti@ccn.yamanashi.ac.jp)}@ccn.yamanashi.ac.jp, <sup>††</sup>[yuu@alps.cs.yamanashi.ac.jp](mailto:yuu@alps.cs.yamanashi.ac.jp),  
<sup>†††</sup>[nakagawa@slp.ics.tut.ac.jp](mailto:nakagawa@slp.ics.tut.ac.jp)

あらまし 本稿では、インターネット上の WEB 文書を利用した音声認識の誤り訂正方法について述べる。特にニュース音声を対象とし、固有名詞に限った誤り訂正実験を行った。正解・誤認識箇所の検出には、2つの音声認識システムを用いる。正解と判定された単語を検索エンジン用のキーワードを選定し、検索により得られた文書中から誤認識単語に対する正解候補を得る。そして、音韻的なマッチング手法により置換判定を行う。実験の結果、誤認識を起こした固有名詞の約 16%を正しい単語に訂正することができた。

キーワード 音声認識誤り, 固有名詞, 誤り訂正, WEB 文書

## An Error Correction Method of Continuous Speech Recognition Using WEB Documents

Hiromitsu NISHIZAKI<sup>†</sup>, Yusuke ITOH<sup>††</sup>, Yoshihiro SEKIGUCHI<sup>†</sup>, and Seiichi NAKAGAWA<sup>†††</sup>

<sup>†</sup> Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi

<sup>††</sup> Dept. of Computer and Media Engineering, University of Yamanashi

4-3-11 Takeda, Kofu, Yamanashi 400-8511 Japan

<sup>†††</sup> Information and Computer Sciences, Toyohashi University of Technology

1-1 Hibarigaoka, Tempaku-cho, Toyohashi, Aichi, Japan 441-8580

E-mail: <sup>†</sup>{[nisizaki,sekiguti](mailto:nisizaki,sekiguti@ccn.yamanashi.ac.jp)}@ccn.yamanashi.ac.jp, <sup>††</sup>[yuu@alps.cs.yamanashi.ac.jp](mailto:yuu@alps.cs.yamanashi.ac.jp),  
<sup>†††</sup>[nakagawa@slp.ics.tut.ac.jp](mailto:nakagawa@slp.ics.tut.ac.jp)

**Abstract** This paper describes an error correction method of continuous speech recognition using WEB documents. We performed an experiment of error correction for news speech automatically transcribed, where we focused on especially proper nouns. Two LVCSR systems were used to detect correctly and incorrectly recognized words. Keywords for the Internet search engine were selected among the correctly transcribed words, then correct candidates for the mis-recognized words were obtained in retrieved documents. Verification between the candidates and the mis-recognized words was utilized DP matching technique using a confusion matrix. In an experimental result, we confirmed that about 16% of mis-recognized proper nouns were corrected into original words.

**Key words** mis-recognized word, proper noun, correction, WEB documents

## 1. はじめに

マルチメディアデータの普及により、そそれのデータを検索する技術の開発が進んでいる。一般的にマルチメディア文書の検索を行うためには、データに対し何らかのインデキシングを行わなければならない。自動的にインデキシングを行う方法として音声認識結果の利用が考えられる。しかし、自動でインデックスを付与するために音声認識技術を用いると、未知語や誤認識の問題に直面する。名詞、特に固有名詞など文書の特異性を表す重要単語は未知語になりやすく、音声認識での正確な書き起こしが難しいとされている。特にニュースなどの日々新しい言葉が出現する時事文書を扱う場合はなおさらである。

そこで、本稿では音声認識結果中の認識誤りを自動的に特定し、訂正するための方法を提案する。認識誤り箇所への訂正の問題は、

- (1) 音声認識誤り箇所の検出
- (2) 検出した誤り単語に対する正解代替候補の作成
- (3) 最適な正解候補の選択

の大きく3つの処理が必要である。

音声認識誤りの訂正方法の研究はいくつか行われている [1] [2] [3] [4]。尾上ら [1] は、ニュース音声の認識誤りを訂正する際に、アナウンサーが読み上げる原稿には、類似した記者原稿があることに着目し、DP マッチングを用いてアナウンサーが読んだニュースに内容に近い記者原稿を探し出し、それと置き換えることで、訂正を行っている。石川ら [2] の方法では、まず、音声認識結果の妥当性を判断するために認識結果を構文解析し、意味的距離を用いて構文木単位での訂正候補を決定する。そして音韻的な特性が類似している用例を検索し、置換を行う。置換処理の妥当性を意味的距離を用いて判断している。沖本ら [3] は認識誤りは分かっていると仮定した上で、認識タスクと同じ内容の文章集合から用例を検索し、文単位での置換処理を行っている。山口ら [4] は、誤り訂正箇所への判別に N-gram と機械学習の手法を用いており、その効果を示している。これらの文献で提案されている訂正手法で共通して言えることとしては、認識する音声と全く同じ、もしくは類似したタスクのコーパス等を利用していることである。こういったコーパス等は、音声認識の言語モデルの学習に用いられることが多い。

そこで本稿で提案する音声認識結果の誤り訂正では、音声認識を行うタスクと全く異なる文書として、近年膨大に増大しているインターネット上の WEB 文書を用いる。WEB 文書を用いる利点としては、1) 膨大な情報 (ニュース記事等) が存在し、それらを手軽に利用できる、2) 音声認識タスクに応じた訂正用の用例コーパス等を用意する必要がない、といっ

たことが挙げられる。また、誤り箇所の検出には、2種類の音声認識システムを用いる [5]。2つのシステムで共通して同じ単語に認識された部分は正解と見なし、一致しない単語を訂正必要箇所と判断し、認識誤り箇所検出を行なう手法を採用する。

本稿で報告する実験では、特に、マルチメディアデータを検索する場合に必要なインデキシングの作成に、音声認識を利用するという点に着目し、インデックス単語の中で特に重要で、誤認識されやすい (未知語になりやすい) 固有名詞に着目し、固有名詞の認識誤りを訂正する実験を行った。今回は、特にニュース動画 (音声) に焦点を当てて実験を行ったが、例えば、講演音声や一般的なテレビ番組、国会答弁などでも応用は可能だと考えている。

2つの音声認識システムを使った誤認識の検出実験では、誤認識した固有名詞の約 90% を誤りであると検出することができた。さらに、自動 WEB 文書検索と DP マッチングによる訂正実験を行った結果、訂正対象の固有名詞の約 41% がインターネット上から自動で検索でき、DP マッチングによる置換判定処理により、約 16% 固有名詞が正しい固有名詞に置換できた。

## 2. 認識誤り訂正処理

### 2.1 処理の概要

処理の流れの概要図を図 1 に示す。

本提案手法では、認識結果 1 文ずつを別々に処理するのではなく、ある一つの塊り (今回はニュース文書を扱うのでその文書) を単位として訂正を行う。まず、2つの音声認識システムからの認識結果に対して、誤り訂正箇所の検出を行う。認識結果に対して訂正必要箇所を判別しておく。同時に、正しく書き起こされたと判断した単語から、WEB 文書を検索するのに用いる単語を選択する。その単語集合を利用して WEB 文書を検索し、得られた WEB 文書集合から誤認識箇所に対する代替単語の候補を選択し、訂正必要箇所と音韻的なマッチングを行い、マッチングした代替単語を訂正必要箇所と置き換える。

以下、訂正処理の詳細について述べる。処理は、

- (1) 音声認識誤り箇所の検出
- (2) 正解代替候補単語の抽出
- (3) 正解代替候補との置換判定処理 (音韻系列同士のマッチング)

の3つに大別される。

### 2.2 誤認識箇所の検出処理

誤認識箇所の検出には、2種類の音声認識システム (SPOJUS [6], Julius3.2p2 [7]) を用いる。NHK のニュース音声データに対して、2つの音声認識システムで共通して同じ単語に認識されると、高確率 (約 95%) で正解であると判定できる [5] ことが実験的に

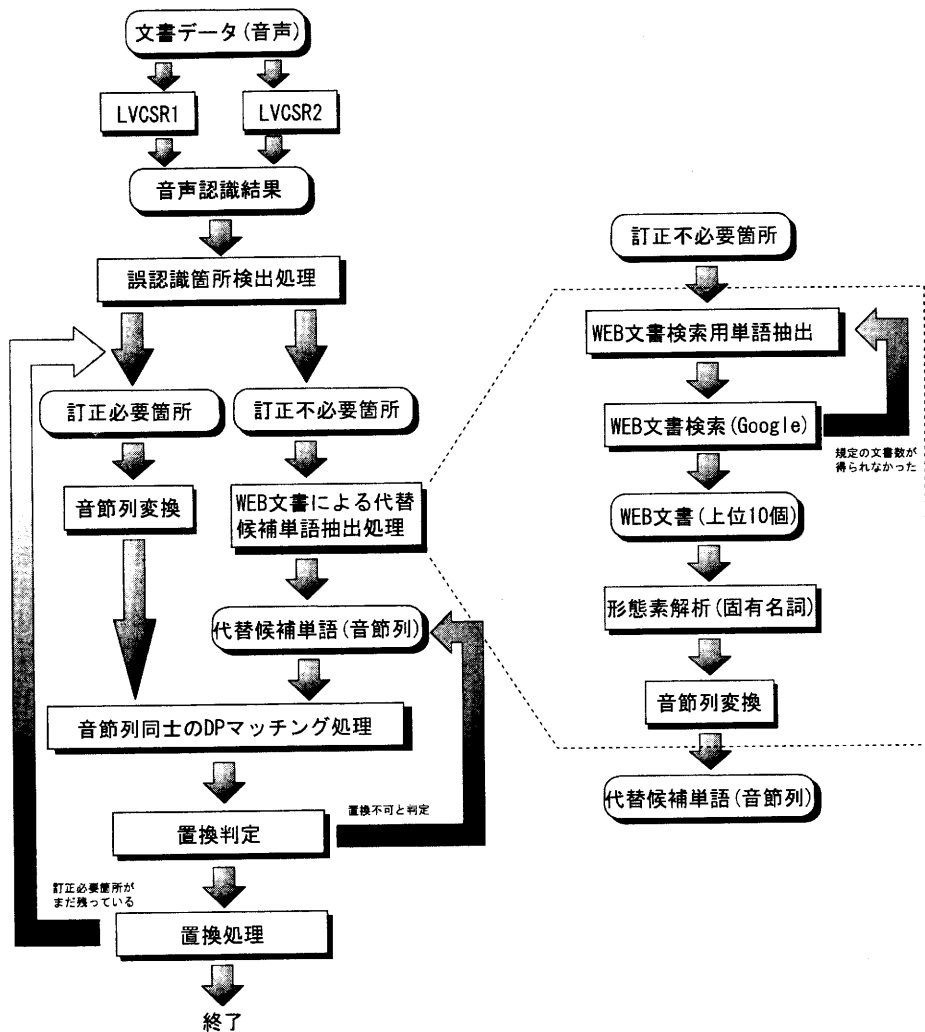


図1 認識誤り訂正処理の流れ

実証されており、図2に示すように2種類の音声認識結果が一致しない単語を訂正必要箇所<sup>(注1)</sup>と判断する。

### 2.3 代替候補単語の抽出

代替候補単語抽出の流れは、図1において破線で囲まれた部分である。認識誤り箇所の代替候補単語の探索に用いる範囲として、インターネット上のWEB文書を用いる。認識結果文と類似した内容の文書集合をインターネット検索エンジン Google<sup>(注2)</sup> (以下“Google”と表記する)を用いて検索する。

Googleで検索するとき用いる検索キーワードと

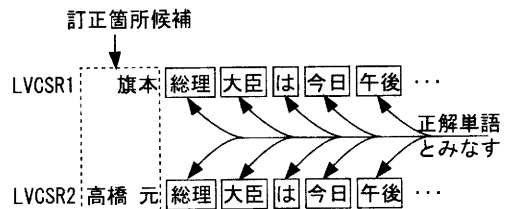


図2 訂正必要箇所の決定

して、訂正必要箇所を含む同じ記事(文書)内から選択した固有名詞を用いる。その理由としては、固有名詞は検索に重要な単語であるためである。当然、検索キーワードに使う固有名詞は2つの認識システムに共通で認識した正解(らしき)単語である。Googleでは検索キーワードは最大で10個までしか受け付け

(注1)：本稿では認識結果に対する誤認識(と判断した)箇所を“訂正必要箇所”、誤認識単語と置き換えるべき単語の候補を“代替候補”と記す。

(注2)：http://www.google.co.jp/

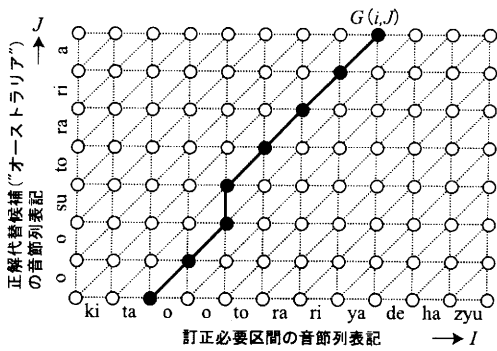


図3 音節列同士の DP マッチングの概念図

ないため、記事の先頭から最初の10個の固有名詞を用いる。Googleで得られた上位10文書を形態素解析器 ChaSen を使って形態素解析を行い、固有名詞のみを音節列に変換し、代替候補の単語集合を作成する<sup>(注3)</sup>。もし、WEB文書が1つも検索されなければ、キーワード集合から固有名詞を1つ取り除き検索を行う。また、10文書未満であれば、検索された文書だけを用いる。

## 2.4 正解代替候補との置換判定処理

### 2.4.1 訂正箇所と代替候補とのマッチング

2.2節で述べた処理により検出された誤認識箇所に対して、代替候補単語と音韻的な距離尺度に基づきマッチングを行い、置換の妥当性を判断する。つまり、訂正必要箇所に対して正解代替候補が含まれているかどうかをワードスポッティングする。

マッチングは音節列レベルの DP マッチングを行う [8]。音節列同士の端点フリーな DP マッチングの概念図を図3に示す。訂正必要箇所の音節列表記を  $X = \{x_1, x_2, \dots, x_I\}$  ( $I$ は訂正必要箇所の音節の数) とし、正解代替候補の音節列表記を  $Y = \{y_1, y_2, \dots, y_J\}$  ( $J$ は代替候補の音節数) と表すこととする。DP マッチングのスコアは、音節同士のコンフュージョンマトリックス  $\{C(r, h)\}$  に基づいて計算を行う。つまり、ある音節  $r$  が与えられたとき音節  $h$  として認識される確率を  $P(r, h) = C(r, h) / \sum_{k \in h} C(r, k)$  とし ( $\phi$ は null 記号),  $I \times J$  の DP マトリックスに対して、格子点  $(i, j)$  での DP スコアを次のように定義する。

$$G(i, j) = \begin{cases} 1.0 & (i \geq 0, j = 0) \\ G(0, j - 1) \cdot P(y_j, \phi) & (i = 0, j > 0) \\ \max \begin{cases} G(i - 1, j) \cdot P(\phi, x_i) \\ G(i - 1, j - 1) \cdot P(y_j, x_i) \\ G(i, j - 1) \cdot P(y_j, \phi) \end{cases} & (i > 0, j > 0) \end{cases}$$

(注3)：代替候補の抽出に関しては文書の検索順位を反映させたスコア等是用していない

ここで、 $P(y_j, \phi)$  は、キーワードを構成する音節の一つ  $y_j$  が脱落する確率を、 $P(\phi, x_i)$  は音節  $x_i$  がキーワード中に挿入される確率をそれぞれ表わしている。 $G(i, j)$  には格子点  $(i, j)$  までに到達する DP パスのうちで最大のスコアを持つパスのスコアが保存される。初期値として、 $G(i, j)$  に  $1.0 (i \geq 0, j = 0)$  を与えることで端点フリーなマッチング (ワードスポッティング) を実現している。 $\max_i G(i, J)$  の正規化スコア<sup>(注4)</sup>が設定閾値よりも大きくなった場合、一致したと判断する。

今回の実験では、訂正必要箇所の音節列表記は、Julius の認識結果から作成している (認識された単語を音節列に変換)。

### 2.4.2 置換判定

訂正必要箇所に対して代替候補でワードスポッティング処理を行い、その結果ある設定閾値以上の値が得られれば、その訂正必要箇所にスポッティングを行った代替候補が含まれていると判断する。閾値を下回った場合には、その候補は訂正必要箇所に含まれていないと判断し、次の代替候補とマッチングを行う。この処理を訂正必要箇所すべてに行う。

## 3. 訂正実験

### 3.1 評価用音声データと認識システム

評価用音声データとしては、NHK ニュース音声データベースから、1996年6月1日の「ニュース7」と「おはよう日本」の収録音声を用いた。記事 (文書) 数は18文書、文数にして175文、6840単語である。音声データにはノイズ (バックグラウンドミュージック、紙をめくる音など) が混入されている。発声者の内訳としてはアナウンサー4名とその他数人のレポーターが含まれており、対話的な発話、レポーターによるノイズ下でのフィールドレポートなどの認識が難しい音声も多数含まれている。ニュース音声の書き起こしに使った音声認識システムとしては、IPA「日本語ディクテーション基本ソフトウェアの開発」プロジェクトから配布されている Julius [7] と、我々の研究室で開発した SPOJUS の2種類を用いる。各音声認識システムで扱う特徴量、音響モデル、言語モデルをまとめたものを表1に示す。両システムとも、音響モデルは、日本音響学会の音素バランス文からなる研究用連続音声認識データベース (ASJ-PB)、および、新聞読み上げ音声コーパス (ASJ-JNAS) を用いて学習を行っている。言語モデルとしては、約5年分のNHK汎用ニュース原稿から学習した語彙サイズ2万の単語 bigram と trigram を用いた。どちらのデコーダも、2パス探索により認識を行い、1パス目では単語 bigram を、2パス目

(注4)： $G(i, J)$  を通過したパスの数で正規化しておく

表 2 各認識システムの単語認識率 [%]

LVCSR	正解率	正解精度
Julius	68.9	63.6
SPOJUS	71.8	56.1

では単語 trigram を、それぞれ使用する。

各認識システムの単語認識率を表 2 に示す。

### 3.2 誤り箇所検出実験

本稿で提案した、2 種類の音声認識システムの結果から、一致しない単語（共通単語が現れるまでの集合体）を訂正必要箇所と判断する。

本実験では固有名詞に限り、2 種類の音声認識システムを用いる事により、どの程度認識誤り箇所を特定できるかを評価するため、誤り箇所検出実験を行なった。またこの方法によって、訂正が必要だと思われる固有名詞が自動検出した訂正必要箇所ほどの程度存在するかを調べた。

表 3 に検出結果を示す。2 つの認識システムで共通に得られた単語は、約 95% の正解率であった。自動検出した訂正必要区間の数（2 つのシステムで共通に認識されなかった区間）は、全部で 775 箇所あり、そのうち訂正の対象となる固有名詞が含まれている箇所は 130 箇所であった<sup>(注5)</sup>。1 つの訂正区間に対して複数の固有名詞が含まれることがあるため<sup>(注6)</sup>、訂正が必要な固有名詞の数は 146 単語である。認識結果を人間が判断し、訂正が必要な固有名詞は全部で 162 単語存在するが、このうち 16 単語に関しては、2 つの認識システムで共通した異なる単語に置換されてしまったので、誤りの自動検出ができなかったことになる。つまり、固有名詞の誤りの自動検出に関しては、この方法を行うことで、誤認識を起こした固有名詞の 90.1% (146/162 = 0.901) を自動で検出できたことになる。なお、1 音素の固有名詞で重要なものはほとんど無いため、音素数 1 の訂正必要箇所は訂正必要箇所を含めていない。また、片方の認識システムで正しく認識された固有名詞もいくらか存在するが、本実験では誤認識として扱う。2 つの認識システム両方で正しく認識されなければ正解としていない。

2 つの認識システムの共通部分から、手動により Google 用の検索キーワード集合を作成し、訂正が必要な 146 個の固有名詞を含んでいる WEB 文書がうまくインターネット上から検索されるかを試みた。その結果、139 個 95.2% (139/146 = 0.952) の固有名詞

(注5)：今回は、訂正必要箇所すべてに固有名詞が含まれていると判断しマッチングを行っている。本来なら訂正区間に対し、固有名詞が含まれるかどうかの判定が必要であると思われる。

(注6)：今回は 1 つの訂正区間に対し、最大で 2 単語の固有名詞を含んでいた。

表 3 誤り箇所検出実験結果

共通部分の単語正解率	94.5%
自動検出した訂正区間	775 箇所
訂正必要固有名詞を含む区間	130 箇所
訂正必要な全固有名詞の数	162 単語
訂正区間に含まれる誤認識固有名詞	146 単語
共通で誤認識された固有名詞	16 単語

表 4 訂正実験結果

訂正単語	146 単語
WEB 文書に含まれる正解単語	76 単語
正しく置換できた単語数	23 単語

に対して<sup>(注7)</sup>、正解（共通）部分の単語を検索キーワードとすることで正しい訂正候補を含んだ単語を検索出来ることが分かった。つまりこの時点で全体の 85.8% の固有名詞が WEB 文書を用いて訂正可能である。

### 3.3 訂正実験

本稿で提案する一連の訂正手法により、どの程度自動で誤りを訂正できるかを確かめるため、実験システムを作成し、訂正実験を行なった。

音声認識結果中の正解（共通で認識された）単語の固有名詞を検索キーワードとし、それによって検索された WEB 文書データに含まれる固有名詞（代替候補）と訂正必要箇所を音節列レベルでの DP マッチング（ワードスポッティング）を行ない、マッチングによって得られたスコア最大の代替候補単語を置換単語とする。

表 4 に訂正実験結果を示す。訂正対象となる固有名詞 146 単語に対し、自動で WEB 上から検索してきた文書中に存在する訂正対象の固有名詞 146 単語に対して、正しく置き換わるべき固有名詞の単語数は 76 単語であった。すなわち、現訂正システムにおいては訂正必要な固有名詞の 52.1% (76/146 = 0.521) に関して自動的にそれを含んだ文書を検索してきてることができたことになる。DP マッチングの最大スコアの代替候補（置換単語）において、正しく置換された単語数は 23 単語であった。すなわち、訂正区間に含まれる誤認識した固有名詞の 15.8% (23/146=0.158) を自動で置換できたことになる。

今回は音声認識率が比較的低い（約 70%）データを対象としたこともあり、全体の訂正率が 16% 弱に留まった。また、本訂正システムでは、1 つの訂正必要箇所において最大 1 単語しか置換することができない。そこで 1 つの訂正必要箇所においても複数の単語を一定の閾値を設け置換を行なうことを可能にさせることにより、一層の自動訂正率の向上が計れ

(注7)：手動で検索作業を行っているのも、もっと良いキーワードセットを用いるとさらにこの数値は上がるかもしれない。

表 1 大語彙連続音声認識の仕様  
(a) Julius

音響モデル	5 状態 3 ループ, 性別依存 (男性), 対角共分散, 16kHz サンプリング, 25ms ハミング窓, フレーム周期 10ms, 音素 triphone モデル (16 混合, 総状態数 3000)
特徴ベクトル	MFCC (12 次元) + $\Delta$ MFCC + $\Delta$ POW (計 25 次元)
言語モデル	NHK 汎用ニュース原稿 (5 年分, 約 12 万文), 語彙サイズ 20000, bigram, trigram モデル

(b) SPOJUS

音響モデル	5 状態 4 ループ, 性別依存 (男性), 全分散, 継続時間制御, 16kHz サンプリング, 25ms ハミング窓, フレーム周期 10ms, 音節モデル (4 混合, 総状態数 1856)
特徴ベクトル	MFCC-seg (12 次元 $\times$ 4 フレームを KL 展開で 24 次元に圧縮 + $\Delta$ MFCC + $\Delta$ $\Delta$ MFCC + $\Delta$ POW + $\Delta$ $\Delta$ POW (計 50 次元)
言語モデル	NHK 汎用ニュース原稿 (5 年分, 約 12 万文), 語彙サイズ 20000, bigram, trigram モデル

ると考える。また、Google で検索を行なう際の検索語を固有名詞のみとしている。人間が検索キーワードを選んで Google で検索すると、約 95% の誤認識固有名詞に対して正解候補を含んだ文書を検索できるが、自動で検索語を選んでもそれが半分程度になった。より文書検索の効率の良い検索語の選定が必要である。

#### 4. おわりに

本稿では、インターネット上の WEB 文書を利用した音声認識の誤り訂正方法について述べた。特にニュース音声を対象とし、検索に重要である固有名詞に限った誤り訂正実験を行った。

2 つの音声認識システムを使った誤認識の検出実験では、誤認識した固有名詞の約 90% を誤りであると検出することができた。そのうちの約 95% の単語に対しては、正解候補の単語がインターネット上の WEB 文書中に存在することが、予備実験により分かった。人間の手によらず自動で検索語を選定し検索を行うと、訂正対象の固有名詞の約 52% がインターネット上から自動で検索でき、さらに DP マッチングによる置換判定処理により、正しい固有名詞に置換できた割合は約 16% であった。訂正必要区間すべてに固有名詞が含まれていると判断しマッチングを行っているが、本来なら訂正区間に対し固有名詞が含まれるかどうかの判定が必要であると思われる。さらに、本来なら訂正必要区間に対しては連続音節認識結果を用いることが望ましい [8]。また、今回の実験は、訂正必要箇所と固有名詞候補とのマッチングは音節系列同士で行ったが、訂正必要箇所の音声と HMM を用いて直接マッチングした方が性能が良いので [8]、この方法と比較する必要がある。

今後は、前述した問題点を解決するような方法を模索し、比較的新しいデータを用いた実験により、提案手法の効果を調べることで、ニュース以外のデータ、例えば、国会答弁などの音声、講演音声、ドキュメンタリー等のニュース以外のテレビ番組の認識結

果の訂正を試み、本手法の効果を確かめたい。また、意味的な情報、例えば、『○○ 総理』という単語列では、○にはほぼ必ず人名が入るといった情報を利用することで、さらに訂正精度の向上を図ることが可能であると考えられる。

#### 5. 謝 辞

この研究では、NHK 放送技術研究所のニュース音声データベース、ニューステキストデータベースを使わせていただいた。これらのデータベースを提供された NHK 放送技術研究所の関係諸氏に感謝致します。また、本研究は平成 15 年度末踏ソフトウェア創造事業の支援を受けています。

#### 文 献

- [1] 尾上和穂, 今井亨, and 安藤彰男. 記者原稿を用いたニュースの音声認識結果の修正法. 春季講演論文集, 1-6-6, pp.11-12. 日本音響学会, 1998.
- [2] 石川 開 and 隅田英一郎. テキストデータを使った音声認識誤りの訂正. 自然言語処理, 7(4), pp.205-228, 10 2000.
- [3] 沖本 純幸, 山本 博史, 隅田 英一郎, and 菊井 玄一郎. 旅行会話基本コーパスを用いた認識誤り訂正の検討”. 情報処理学会研究報告, 2001-SLP-39, pp.159-164, 2001.
- [4] 山口辰彦, 酒向慎司, 山本博史, and 菊井玄一郎. 信頼度尺度に基づく音声認識誤りの検出および誤り訂正. 秋季講演論文集, 3-Q-14, pp. 153-154. 日本音響学会, 2003.
- [5] 宇津呂武仁, 西崎博光, 小玉康広, and 中川聖一. 複数の大語彙連続音声認識モデルの出力の共通部分を用いた高信頼度部分の推定. 電子情報通信学会論文誌 D-II, Vol.J86-D-II, No.7, pp.974-987, 2003.
- [6] 甲斐彦彦, 廣瀬良文, and 中川聖一. 単語 n-gram 言語モデルを用いた音声認識システムにおける未知語・冗長語の処理. 情報処理学会論文誌, Vol.40, No.4, pp.1383-1394, 2000.
- [7] 河原 達也, 李 晃伸, 小林 哲則, 武田 一哉, 峯松 信明, 嵯峨山 茂樹, 伊藤 克互, 伊藤 彰則, 山本 幹雄, 山田 篤, 宇津呂 武仁, and 鹿野 清宏. 日本語ディクテーション基本ソフトウェア (99 年度版) の性能評価. 情報処理学会研究報告, 2000-SLP-31-1, pp.9-16, 2000.
- [8] 西崎博光 and 中川聖一. 音声認識誤りと未知語に頑健な音声文書検索手法”. 電子情報通信学会論文誌. 電子情報通信学会論文誌 D-II, Vol.J86-D-II, No.10, pp.1369-1381, 2003.