

話題による音声認識誤り単語の補正手法 Compensation method for speech recognition error word based on topic sentence

角地 良太[†] 芋野 美紗子[†] 土屋 誠司[‡] 渡部 広一[‡]
Ryota Kakuchi Misako Imono Seiji Tsuchiya Hirokazu Watabe

1. はじめに

音声認識の分野ではこれまでに様々な研究が行われており、近年では人間との会話が可能でロボットが注目されている。しかしそれらを実現するためには話者や環境、語彙などの変化によって認識率が左右されやすいという問題を解決する必要がある。人間はこのような音声が入り込んで聞き取れない雑音環境下においても会話を成立させることができる。これは、人間が1音1音のすべてを認識しているのではなく、聞き取れた言葉から内容を理解し、聞き取れなかった言葉を補正して認識しているからである。これを音声理解と呼ぶ。そこで本稿では、音声理解の立場から音声認識装置 Julius の認識処理結果に対して文の意味・内容から話題を推測し、誤認識した単語をその話題とより関係の深い単語へ補正する手法を提案する。

2. 関連技術

2.1 連続音声認識装置 Julius^[1]

音声認識装置 Julius (以下 JU と呼ぶ) とは、単語辞書の登録語彙数が 6 万語の不特定話者用音声認識装置であり、音響モデル及び言語モデルから成る。

入力：優勝決定戦では同点の場合 PK 戦を行う

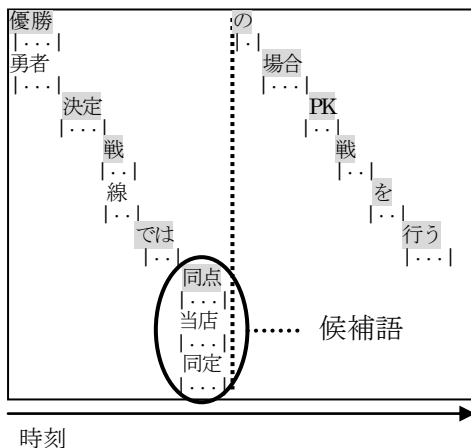


図 1：各音声区間における音声認識処理情報

図 1 に JU の単語認識処理情報を示す。図 1 の右軸を時刻とする。JU の音声認識処理は、認識処理をした各単語に対して候補語(図 1 における丸に囲まれた単語の集まり)をもっており、そこから正解とする単語(以下正解語と呼ぶ)を選出している。

2.2 記事分類方式^[2]

記事分類方式とは、Web から抽出したカテゴリ属性と記事関連度計算方式を用いて記事を政治・経済・社会・スポーツの 4 つのカテゴリ(話題)に分類する手法である。属性とは、意味特徴を表す単語の事であり、カテゴリについて表したものをカテゴリ属性と呼ぶ。記事関連度計算方式とは、記事を 1 つの概念、記事の索引語を属性と見なし、索引語の出現度合と特定性を表す尺度に基づいて決めた重要度(重みと呼ぶ)によって記事 A と記事 B との関連の深さを定量的に表す計算方式のことである。

3. 音声認識誤り単語の補正手法

音声認識装置によって誤って認識された文の認識処理結果を見てみると、誤って認識した単語(以下誤認識単語と呼ぶ)にはいくつかの候補となる単語(以下候補語と呼ぶ)が存在しているのがわかる。さらにその候補語を確認すると本来選出されるべきである単語が存在している場合が多い。例として、次の JU の認識処理結果を示す。

【音声入力】優勝決定戦で同点の場合、PK 戦を行う。

【認識処理結果】優勝決定戦で当店の場合、PK 戦を行う。

認識処理結果は音声入力に対して、「同点」を「当店」と誤って認識している。人間であれば、文を構成する「優勝決定戦」や「PK 戦」という単語からこの文が「スポーツ」や「サッカー」に関する内容であると連想することができる。また「スポーツ」や「サッカー」から特徴的な単語を連想することもできる。その連想によって人間は「当店」ではなく「同点」であろうと判断することができる。このような連想機能を音声認識の処理に加えることで有効な音声理解が行えると考える。そのことを踏まえ、音声認識の認識処理結果を話題を用いることで誤認識単語を検出し、正解語に補正する手法を提案する。これを用いて作成したシステムが音声言語理解システムである。

4. 音声言語理解システムの流れ

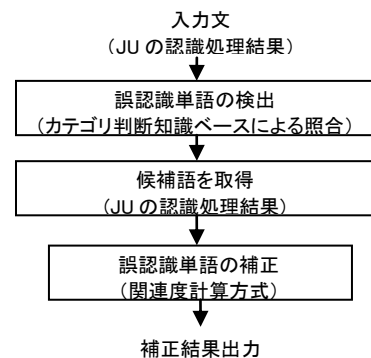


図 2 音声言語理解システムの流れ

[†] 同志社大学大学院工学研究科

Graduate School of Engineering, Doshisha University

[‡] 同志社大学理工学部

Faculty of Science and Engineering, Doshisha University

図2に音声言語理解システムの流れを示す。JUの認識処理結果(入力文)をシステムに入力し、誤認識単語を検出する。次に検出された誤認識単語の候補語を認識処理結果より取得し、システムによって正しいと判断された候補語を誤認識単語と置換し補正結果として出力するのがシステムの一連の流れである。

4.1 誤認識単語の検出

誤認識単語の検出は、記事分類方式とカテゴリ判断知識ベースによって行う。カテゴリ判断知識ベースとは、記事分類方式で分類する4つのカテゴリそれぞれに対して関連性のある、つまり、「特徴的で不自然でない」名詞、またそれらの意味属性(ノード)を集めた知識ベースのことである。ノードはNTTシソーラス^[4]を用いることで取得する。

まず入力された認識処理結果を記事分類方式によってカテゴリに分類する。次に選ばれたカテゴリに属するカテゴリ判断知識ベースによって認識処理結果の中からカテゴリに不自然な単語を探し出す。カテゴリに不自然な単語が誤認識単語であろうと考えられる。

これが誤認識単語の検出である。

誤認識単語検出の例を挙げる。

【音声入力】団体個人ともメダルのチャンス

【認識処理結果】団体個人ともメタルのチャンス

システムに入力された認識処理結果は記事分類方式によって「スポーツ」のカテゴリに分類される。次に認識処理結果に含まれる名詞のノードをNTTシソーラスによって調べ、それらノードと名詞をカテゴリ判断知識ベースに集約されている単語と一致するかをマッチングによって判断する(以下照合と呼ぶ)。この場合「メタル」という名詞が、カテゴリ判断知識ベースに集約されている単語と一致しないため、「スポーツ」のカテゴリには不自然な単語であると判断される。よってこの場合、誤認識単語は「メタル」になる。

4.2 誤認識単語の補正

誤認識単語の補正には、まず4.1節によって検出した誤認識単語の候補語を認識処理結果より取得する。次に取得した候補語が認識処理結果の分類されているカテゴリに対して関連性のある単語かを調べるためにカテゴリ判断知識ベースと照合する。そして誤認識単語であると判断された単語と候補語を置換して文を作成し、カテゴリとの記事関連度計算を行う。その結果、関連度の最も高くなった文を訂正文として出力する。

誤認識単語の補正について例を挙げる。4.1節の例より「メタル」が誤認識単語と判断されるため、その候補語である「メダル」、「メタン」を認識処理結果より取得する。次にこれら候補語が「スポーツ」のカテゴリに対して関連性のある単語かを判断するためにカテゴリ判断知識ベースと照合する。その結果、「メタン」という単語は「スポーツ」のカテゴリには不自然な単語と判断されるために候補語にはならず、不自然でない単語と判断された「メダル」という単語は候補語となる。そして候補語を誤認識単語と置換し文を作成する。

【訂正文候補】団体個人ともメダルのチャンス

しかし、入力した認識処理結果が本来誤認識単語を含まない文であった場合であることも考えられる。そのため、訂正文候補とするのは先ほど作成した文とシステムに入力した認識処理結果の文になる。

【訂正文候補】団体個人ともメダルのチャンス

【認識処理結果】団体個人ともメタルのチャンス

認識処理結果と作成した訂正文候補の2文を「スポーツ」のカテゴリとの関連度を記事関連度計算方式によって調べ、関連度の値が最も高くなった訂正文として出力する。この場合、候補語を取得し作成した訂正文候補が認識処理結果より関連度が高くなったため、訂正文候補が訂正文となる。

5. 音声言語理解システムの評価

音声言語理解システムの評価を行った。評価方法は、検出精度(誤認識単語を検出できるかの精度)と訂正精度(正解語に訂正できたかの精度)である。表1に結果を示す。

表1 検出精度と訂正精度

	本研究
検出精度	17.1%
訂正精度	15.4%

結果は76文中13文の誤認識単語を正しく検出し、そのうち2文を正しい訂正文に補正できた。検出精度は17.1%、訂正精度は15.4%となった。

また、誤認識単語が含まれない音声出力文に対しても評価した。表2に結果を示す。

表2 正しい文を正しいと判断できるかの精度

	本研究
判断精度	86.3%

結果は219文中189文が、誤認識単語がなく正しい文と判断され、精度は86.3%となった。

6. おわりに

本稿では、人間が普段行っている音声理解に着目し、文の話題を用いて誤認識単語の補正手法の提案を行った。この提案手法を基に構築した音声言語理解システムの評価は、検出精度17.1%、訂正精度は15.4%となった。これらの評価より提案手法の精度向上が望まれる。そのためには、話題決定を担う記事分類方式の精度向上、カテゴリ判断知識ベースの集約をより厳密にすることが求められる。

謝辞

本研究の一部は、科学研究費補助金(若手研究(B)21700241)の補助を受けて行った。

参考文献

- [1]京大大学、情報処理振興事業協会(IPA)2, “大語彙連続音声認識デコーダ Julius”, <http://Julius.sourceforge.jp/>
- [2]藤江 悠五, 渡部 広一, 河岡 司, “Web情報から抽出したカテゴリ属性と記事関連度計算を用いた記事分類方式”, 第21回人工知能学会全国大会論文集 1G3-5, 2007.6/
- [3]NTTコミュニケーション科学研究所, “日本語語彙体系”, 岩波書店, 1997