

音声認識結果の補正支援ツールの開発

安東 徹† 川波 弘道††

† 津山高専専攻科 電子・情報システム工学専攻 †† 津山高専 総合理工学科

1. はじめに

現在では、音声認識を使った技術が多く開発され、世間に広く普及している。しかし、認識精度が完全ではないために、本来の発声内容と異なる音声認識をしてしまう場合がある。そこで、本研究では、認識結果の誤りを容易に修正できる手段を検討した。

先行研究[1]では、音声認識結果のうち不明瞭な単語に対して復唱し、再認識させることで誤りを訂正していた。正確な音声認識結果を提示するにあたり、認識結果を有効活用する方向性で考察を進めた。認識エンジンは、このような単語であるという可能性を示す単語仮説を複数持つ。最終的に出力される音声認識結果は、音声認識辞書に基づき、最も有力な候補が決定される。しかし、その過程で切り捨てられた他の認識結果の候補が、実際には正しい内容だったという場合がある。

本研究の目的は、Julius[2]の音声認識結果の補正を支援するツールを開発することである。様々な認識結果の候補を可視化するユーザインタフェースを提供し、認識結果の誤りを容易に修正できることを目標とした。

2. ツールの仕様検討

図 1 にツールの具体的な処理の流れを示す。

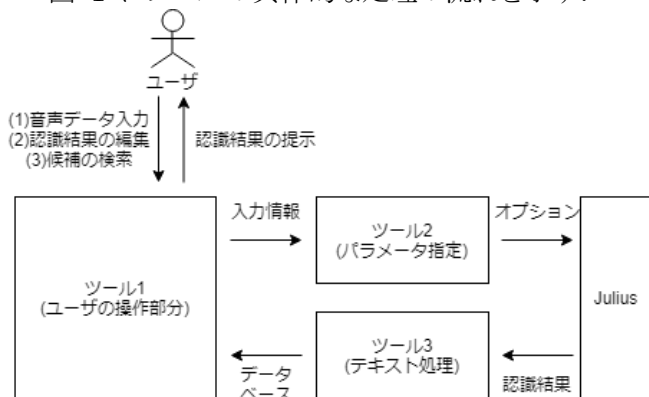


図 1. ツールの内部での動作

ユーザが操作する部分はツール 1 のみとなっている。まず、ユーザはテキストデータへの書き起こしをしたい音声データを指定する。その後、入力された音声データやパラメータの値をツール 2 に渡す。ツール 2 では、渡された値を Julius で実行するコマンドに組み込み、Julius に各種パラメータを指定して音声認識させる。Julius が音声認識を終えると、認識結果が記述されたファイルを出力する。そのファイルをツール 3 でテキスト処理し、各文仮説の単語をまとめたデータベース

を作成する。そして、そのデータベースをもとに認識結果を書き起こし、ユーザに提示する。ユーザは、書き起こされた認識結果に誤りを発見した場合、その内容を編集することができる。また、その単語が他にどのような単語仮説を保有していたかを検索することができる。

3. ツールの特徴

図 2 に本研究で開発したツールの外観を示す。

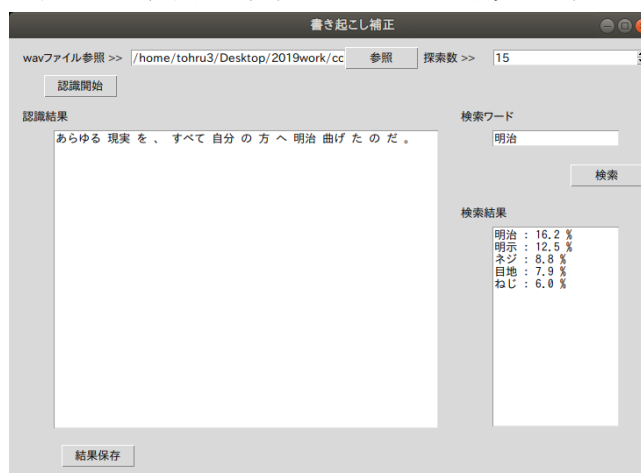


図 2. 開発したツールの動作画面

ツール上部の「参照」ボタンを押し、書き起こしをしたい音声ファイルを選択した後、「認識開始」ボタンを押して音声認識をする。すると、Julius に音声ファイルとパラメータが渡され、「認識結果」に書き起こしが行われる。書き起こされたテキストに誤りがあると考えられる場合、「認識結果」のテキストボックスで編集することができる。この際、誤りの補正を支援する機能として、「検索」ボタンを押下することでほかの候補を検索することができる。検索された単語に対して、他にどのような候補があったか、またその候補の単語信頼度を「検索結果」に出力している。この検索結果を参照しながら、書き起こされた認識結果の誤りを補正することができる。

4. 今後の課題

当ツールの有用性については、評価実験の結果より判断する必要がある。より便利なツールにするためには、機能の追加を行う開発の継続が必要である。

参考文献

- [1]大田健翔, 秋田祐哉, 河原達也: 講演音声認識結果の誤り箇所への復唱入力を用いたノートテイクシステム, 情報処理学会第 77 回全国大会講演論文集, 1 号, pp.125-126(2015).
- [2]大語彙連続音声認識エンジン Julius:<http://julius.osdn.jp/>.