

音声認識を用いた要約筆記補助システムの検討

Examination of Note-taking support system using voice recognition

伊差川 恵悟, 小渡 悟
Keigo ISAGAWA, Satoru ODO
産業情報学部 産業情報学科
Department of Industry and Information Science
沖縄国際大学
Okinawa International University
Email: 12DB010@okiu.ac.jp

あらまし：聴覚障害者に対して支援者が文字を使って情報保障をする「要約筆記」というのがある。1対1でのコミュニケーションにおいて要約筆記の支援が得られない場合でのサポートを目的とし、音声認識機能を用いて、音声を自動で文字化して提示するシステムの構築を試みたので報告する。

キーワード：要約筆記, 聴覚障害者, 情報保障, 音声認識

1. はじめに

「要約筆記」とは、健聴者が話した内容を聴覚障害者に文字で伝える聴覚障害者に対する情報保障の一つである。

「要約筆記」は沖縄国際大学でも聴覚障害者の学生に対して「福祉・ボランティア支援室」を通して「ノートテイク」という形で行われている。

最近では、聴覚障害者に対しての情報保障システムとして、アイセック・ジャパンの「字幕電話」や東京都北区議会で、議員の斉藤 里恵氏が利用している議会での発言を画面に表示させるシステムなどがある。

どちらのシステムも聴覚障害者にとって有用なシステムであるが、聴覚障害者が利用しやすくするためにはより手軽に使える情報保障システムが必要である。

そこで、本研究では音声認識機能を用いて聞き取った音声を文字にして、画面に提示するシステムの構築を試みたので報告する。

2. 提案システム 1 「『w3voice』を利用したシステム」

2.1 提案システム

音声認識システムとして和歌山大学聴覚メディア研究室で開発された Java 版 w3voice システムを用いた。w3voice の API を用いることで音声認識機能をウェブブラウザで利用することが可能となる。

w3voice システムでは音声認識 JavaScript ライブラリ w3voiceIM.js を公開しており、HTML コードに追加することでウェブサイトのテキストボックスに音声認識機能を持たすことができる。利用者と開発者は専用ソフトウェアのインストールが必要なく、Java と JavaScript が使えるウェブブラウザであれば利用可能となっている。

提案システム 1 では音声認識 JavaScript ライブラリ w3voiceIM.js を用いることでウェブブラウザ上

のテキスト入力エリアでダブルクリックすることで w3voice システムの画面が起動するようにした。画面上の「Push here.」ボタンを押し続けることで録音を行う。録音後は、音声認識の結果を確認してから「add」をクリックすることでテキストとして入力される。

実行画面の例を図 1 に示す。利用者が読みやすいように画面構成は PC での要約筆記を行うためのソフトウェアである「IPtalk」や「まあちゃん」の画面と同じようにし、黒背景と大きめの白い文字にした。



図 1 w3voice を使用したシステム

2.2 実験結果

w3voice システムを用いた要約筆記システムを実際にサーバにて動作させた。ウェブブラウザは「Internet Explorer」、「Firefox」、「Google Chrome」にて動作確認を行った。

黒いテキストボックスの中をダブルクリックすると w3voice システムが起動した。録音中はメーターが動き、マウスのボタンを離すまで録音される。また、録音が終わった後は w3voice のサーバに送信され、音声認識した結果が表示される。

被験者として大学生 1 名に利用してもらったところ、ダブルクリックして録音している間マウスをクリックし続ける必要や改行は自動ではされず、音声認識を行うたびに改行する必要がある、操作性に難があると報告があった。また、画面は改行の有無にかかわらず、文字が上から下へ向かうため、「IPtalk」の自動スクロールと逆向きになることから、IPtalk の

利用者から表示画面に戸惑うと報告があった。また、マイクの位置の問題から話した内容とは違う言葉が表示されるなどの不具合も多かった。

3. 提案システム 2 「『MIT App Inventor 2』で開発した、Android スマートフォン用アプリケーション」

3.1 提案システム

スマートフォンやタブレットを使い、マイクで音声を読み取って文章を画面に表示する Android アプリの開発を行った。開発環境として MIT App Inventor 2 を使用した。

App Inventor とは Google が開発し現在は MIT が提供している Android アプリケーション開発システムである。図 2 に開発画面の例を示す。

Android アプリの画面構成は、w3voice を使用した開発システムと同様に、パソコン要約筆記の画面設定で推奨されている黒背景と白文字の構成にした。また、音声を聞かせるためのボタンを下のほうに設置し、音声認識は MIT App Inventor 2 の SpeechRecognizer 機能を利用した。

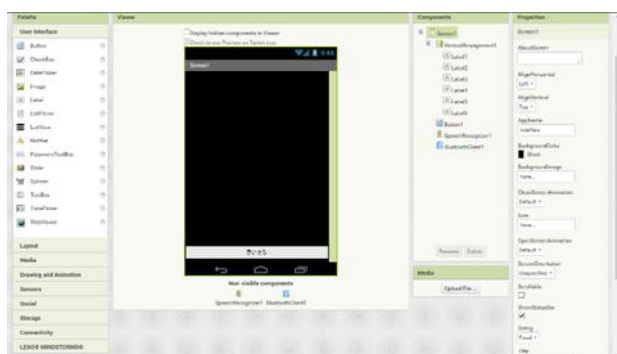


図 2 「MIT App Inventor 2」の画面

画面に表示された文章は編集できないようになっているが、図 3 の画面「きかせる」ボタンの横にある「編」のチェックボックスを入れることによって、文章の訂正が可能である。

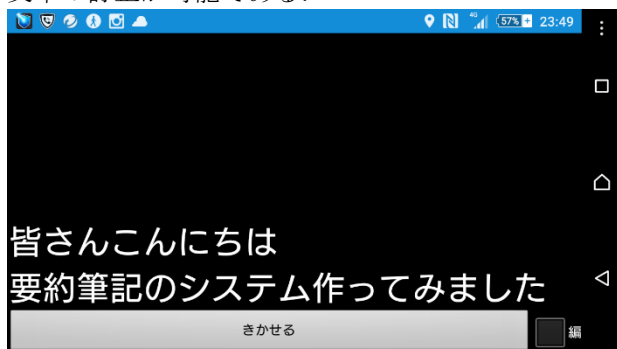


図 3 開発したアプリケーションの画面

3.2 実験結果

Android 5.0 のスマートフォンにて開発アプリケーションが利用可能であるかを実験した。

画面下の「きかせる」ボタンを押すと Google の音声認識システムが起動した。話しかけるとその結果が画面に表示し、表示された文字はパソコン用の要約筆記ソフトと同様に、下から上へスクロールさせて表示が行えた。

4. 考察

w3voice を使用して開発したシステムは、実際に行われているパソコン要約筆記の画面に近づけることができた。しかし、聞き取った音声は自動入力ではなく、また録音をしている間は「Push here.」のボタンをマウスでクリックし続けなければならないなどの問題点が多かった。

MIT App Inventor 2 で開発したシステムは、Android のスマートフォンやタブレット向けのアプリケーションとして開発した。利用する音声認識システムを Google の音声認識に変更し、音声認識の精度は前回と比較して向上した。

音声認識の精度が向上した理由は、パソコンではなくスマートフォンを利用したことにより、利用者が声を機器に近づけて話しやすくなったからだと考えられる。

5. まとめ

本研究では、聴覚障害者に対して、文字を使って情報保障をする「要約筆記」をより身近にするためにシステムを開発した。

w3voice を利用したウェブブラウザで利用する要約筆記システムは、ウェブブラウザで誰でも利用可能であるが、操作性はマイク位置による音声認識の精度が異なることから改善をする必要がある。

MIT App Inventor 2 で開発した Android スマートフォン用アプリケーションは、マイクが聴きとる範囲が狭いため、講義には向かないが、聴覚障害者と 1 対 1 でコミュニケーションをとるときには有用であろう。今後も、文章を要約する機能や、音声認識の結果を Bluetooth で送信しパソコンで表示できるようにする機能などを追加する予定である。

謝辞 本研究は JSPS 科研費 15K00292 の助成を受けたものです。

参考文献

- (1) 山城 秀生 監修, 上村 博一 著:字が話す 目が聞く 日本語と要約筆記(2003)
- (2) 「要約筆記者養成テキスト」作成委員会:厚生労働省カリキュラム準拠 要約筆記者養成テキスト 上 (2013)
- (3) 音声入力・音声認識 Web アプリ w3voice Laboratory <http://w3voice.jp/>
- (4) MIT App Inventor 2 <http://ai2.appinventor.mit.edu/>
- (5) MIT App Inventor2 攻略 WIKI <http://appinventor2.wiki.fc2.com/>