

障害者自立支援機器等開発促進事業

聴覚障害者の日常生活支援機器－
音声認識し文字表示するメガネや携帯可能な支援機器に関する開発

(2年計画の1年目)

平成22年度 報告書

開発機関 株式会社アドバンスト・メディア

平成23(2011)年 4月

聴覚障害者の日常生活支援機器
音声認識し文字表示するメガネや携帯可能な支援機器 概要

代表機関名 株式会社アドバンスト・メディア

※ 全体の概要

聴覚障害者の方々の日常会話を支援することを目的として、聴覚障害者と対話する相手の発話を音声認識し、文字表示する機器を開発する。

日常会話の音声やそれを忠実に書き起したテキストが十分に無いという問題があり、実用精度の日常会話の音声認識はまだ実現できていない現状がある。本年度は音声収集、書き起しによる音声のテキストデータ化、テキストデータの追加学習による音声認識ソフトウェアの開発、および、音声筆談ボードの試作機の開発を行った。来年度は試作機を使ってモニタ試験に臨む。

※ 試作した機器またはシステム

■ 音声認識筆談ボード

日常会話支援機器の試作機器としてスマートフォン(iPad)上で音声認識筆談ボードを開発した。これを用いて音声も収集も行った。

iPad の内蔵マイクで音声認識した文字列を表示し、その上から手書き入力も可能にした。

音声認識が誤認識しても手書き入力で下図のように簡単に訂正文字を表示することができる

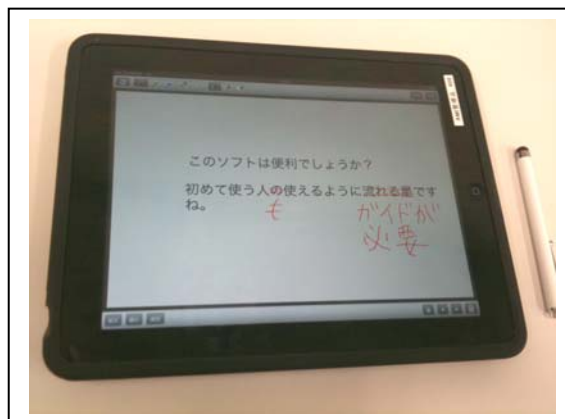


図1 音声認識筆談ボードの試作機の概観

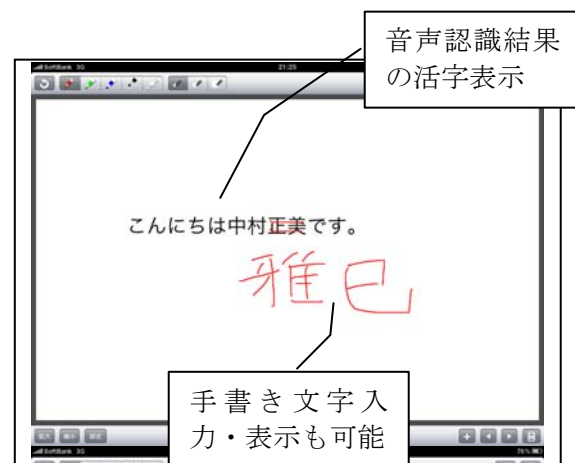


図2 音声認識筆談ボードの表示画面

■音声認識ソフトウェア

弊社が既に保有している、会議音声用の音声認識ソフトウェアのデータベース(辞書および言語モデル)に今回収集した日常会話の音声を書き起したテキストデータを追加学習して、日常会話用の音声認識ソフトウェアを開発した。音声認識のブロック図を図3に示す。今回は難聴者の協力を得て、日常会話の音声を大量に集め、一字一句文字列に書き起しをして、まずは言語モデルおよび辞書を日常会話用にカスタマイズ開発した。既に会議用の話し言葉対応の言語モデルを保有しているのでその言語モデルに追加学習する方法を取った。音響モデルのカスタマイズは来年度実施する予定である。

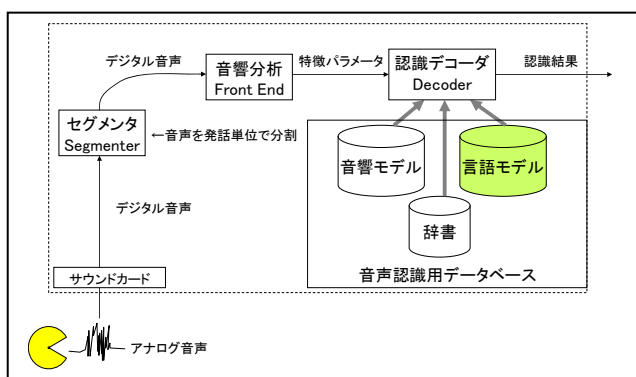


図3 音声認識のブロック図

以上

目 次

I.	報告	-----	1
II.	開発成果の公表に関する一覧表	-----	19
III.	開発成果の公表に関する刊行物・別刷	-----	19

障害者自立支援機器等開発促進事業
総括報告書

聴覚障害者の日常生活支援機器－
音声認識し文字表示するメガネや携帯可能な支援機器に関する開発

開発代表機関 株式会社アドバンスト・メディア

開発要旨

聴覚障害者の方々が日常生活を送る上で困難と感じている、日常会話を支援する手段を開発することが目的である。具体的には、聴覚障害者と対話する相手の発話を音声認識し、文字表示してコミュニケーションをとることを支援する機器を開発する。

本年度は音声収集、書き起しによる音声のテキストデータ化、テキストデータの追加学習による音声認識ソフトウェアの開発、および、音声筆談ボードの試作機の開発を行った。

開発代表者 中村雅巳
(株)アドバンスト・メディア
開発本部医療・議事録ソリューション部
部長

開発者 坂本伸幸
開発者 小黒 玲
開発者 谷 直樹
(株)アドバンスト・メディア
技術本部基礎技術部
研究員

開発者 小澤光広
開発者 青木秀仁
開発者 中村正規
開発者 立川公一
(株)アドバンスト・メディア
開発本部
システムエンジニア

A. 開発目的

一般社会においては、人と人とのリアルタイムなコミュニケーションは音声による会話によって行われることが多く、聴覚障害者にとってはその音声言語がコミュニケーションのバリアとなっている。聴覚障害者の数は約 34 万人(※1)いるとされているが、聴力が衰えた高齢者や「話すのにやや不便を感じる」というレベルの人まで含めると、その十倍以上はいるともいわれ、その膨大な数の人々が日常生活で大変な不便を感じている。

開発者の所属する(株)アドバンスト・メディアでは 13 年前の会社設立当初から、社団法人全日本難聴者・中途失聴者団体連合会(以後、全難聴と略する)と協力し、音声認識の活用を試みてきた。その過程で議事録作成支援システムを開発し製品化した実績がある。昨年からは横浜市中途失聴・難聴者協会(以後、浜難聴と略する)の川井会長から携帯できる音声認識機器の開発の相談を受け、今回の開発のきっかけになった。

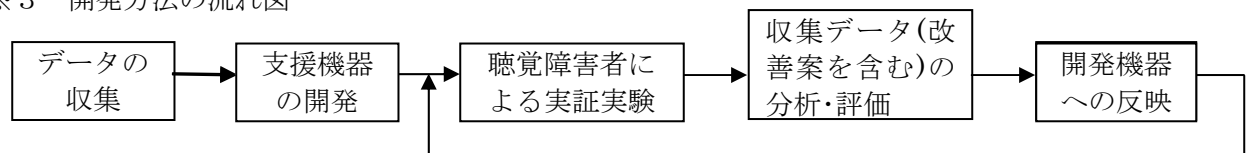
精度の高い音声認識を開発するためには大量の発話テキスト情報から言語モデル・辞書を作成する必要があるが、日常会話に関しては一字一句そのまま文字化したテキストデータは世の中にはほあまり存在していないという問題があった。

そこで、今回は日常会話やお店での応対発話などの音声を幅広く収集し、一字一句書き起しをして開発のための基礎データを作成、その基礎データを元に、音声認識ソフトウェアや支援機器の開発を行った。来年度は実証実験(モニタ実験)を行って評価したい。評価結果は開発物に反映し、実証実験するというサイクルを繰り返して完成度を上げる(※3)。本開発促進事業の終了後はすぐに製品化しその年の秋までにはサービスを開始したい(※4)。

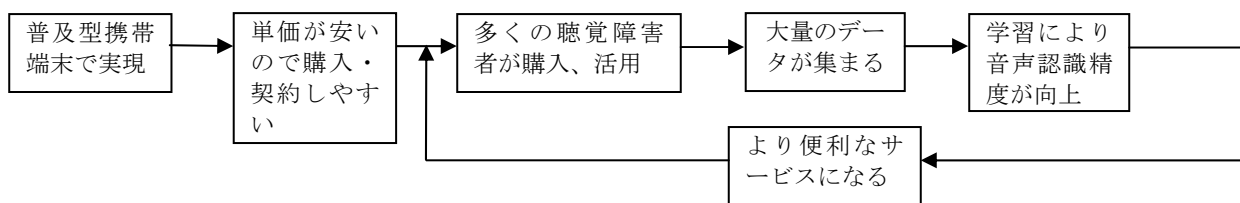
※1.「平成18年身体障害児・者実態調査結果」厚生労働省社会・援護局障害保健福祉部企画課発行

※2.「音声認識の動向 [I] 話し言葉音声認識」, 古井貞熙, 電子情報通信学会誌, Vol.89, No.8, pp.746-751 (2006)

※3 開発方法の流れ図



※4 事業の流れ図



B. 開発する支援機器の想定ユーザ

聴覚障害者および聴力の衰えた高齢者が想定ユーザである。

実際に支援機器に向かって発話するのは、上記想定ユーザと対話する人達である。

C. 開発体制

■開発メンバー

●統括

中村雅巳((株)アドバンス・メディア)

●試作機器の開発

青木秀仁((株)アドバンス・メディア)

中村正規((株)アドバンス・メディア)

小澤光広((株)アドバンス・メディア)

●音声認識ソフトウェアの開発

坂本伸幸((株)アドバンス・メディア)

小黒玲((株)アドバンス・メディア)

谷直樹((株)アドバンス・メディア)

●データ整理

立川公一((株)アドバンス・メディア)

■アドバイザー ※敬称略

鈴木恵子

北里大学医療衛生学部リハビリテーション
学科言語聴覚療法学専攻

講師

■当事者団体(協力者) ※敬称略

横浜市中途失聴・難聴者協会

川井節夫、須山優江、小川由恵、大林
紀夫、長尾重之、永島陽子

東京都中途失聴・難聴者協会

小川光彦、福田能文

一般

渡辺儀一、寺田翔

D. 試作した機器またはシステム

1. 試作機器

日常会話支援機器の試作機器としてスマートフォン(iPad)上で音声認識筆談ボードを開発した。これを用いて音声も収集も行った。

iPad の内蔵マイクで音声認識した文字列を表示し、その上から手書き入力も可能にした。

音声認識が誤認識しても手書き入力で図のように簡単に訂正文字を表示することができる。

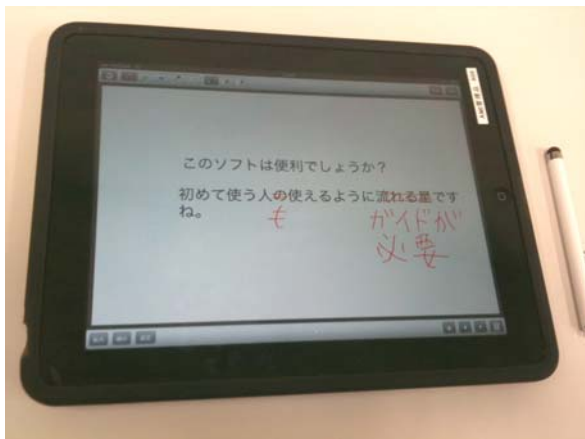


図-1 音声認識筆談ボードの試作機の概観

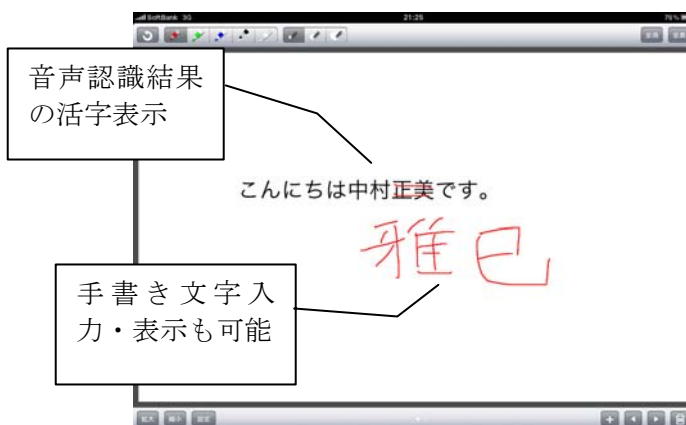


図-2 音声認識筆談ボードの表示画面

【操作方法】

①音声認識筆談ボードを起動

(1) 左上の白い四角のマーク(UbiWriteDSR)を指で押す(タップ)する。



図-3 音声認識筆談ボードのアイコン

(2) アプリを起動すると、右下のような画面が表示される。

背景が白い部分に音声認識や手書きを使って入力した内容が表示される。

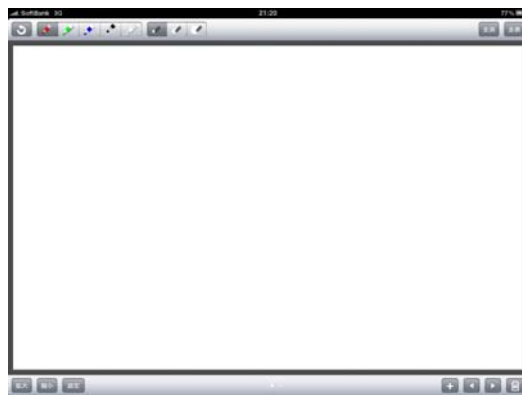


図-4 音声認識筆談ボードの初期画面

②音声認識・録音の開始・終了

(1) 表示画面の任意の場所を押している間、録音されるので、内臓マイクに向かって発話する。

(2) しゃべっている最中は音声の波形が表示される。(音声入力中)

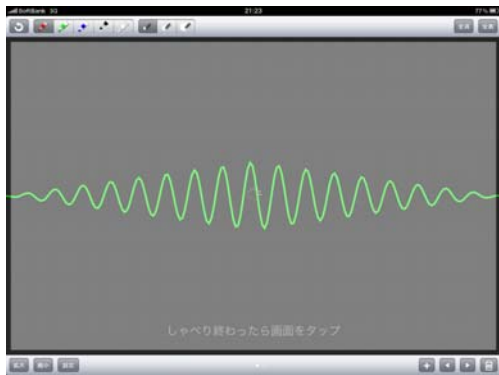


図-5 音声入力中画面表示

(3) しゃべり終わって、画面から指を離すと録音が終了し、音声認識が開始し、認識が終了すると指で指定した位置から文字列が表示される。



図-6 音声認識結果表示

今回は音声収集するのも目的であるため、しゃべった音声は自動的にサーバーに録音される。

③手書き入力

直接、指を画面に当てて手書きで文字を描くことも可能である。

音声認識で変換誤りが発生した時は、再度発話し直すか、変換誤り部分を手書き入力で訂正する。

ペンの太さや色は上部のパレットで変更可能。また、複数枚のシート(ページ)を使用することもできる。

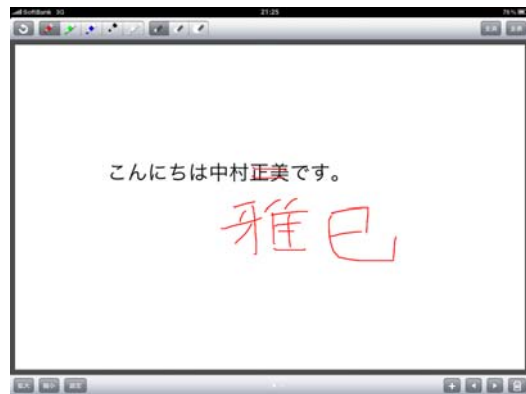
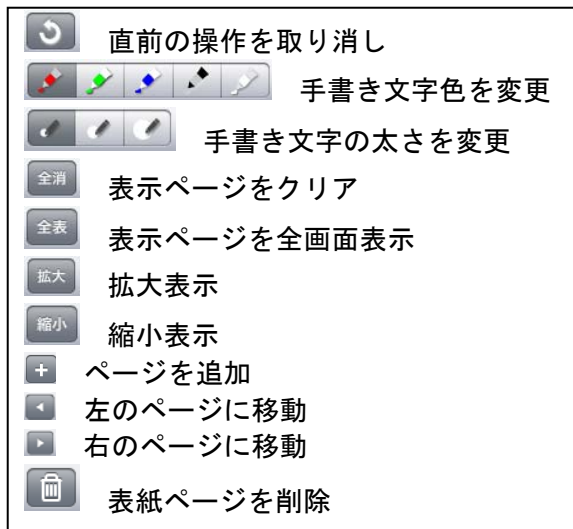


図-7手書き入力結果

④その他の機能

以下のようなボタン機能がある。



2. 音声認識ソフトウェア

弊社が既に保有している、会議音声用の音声認識ソフトウェアのデータベース(辞書および言語モデル)に今回収集した日常会話の音声を書き起したテキストデータを追加学習して、日常会話用の音声認識ソフトウェアを開発した。

音声認識のブロック図を図-8に示す。

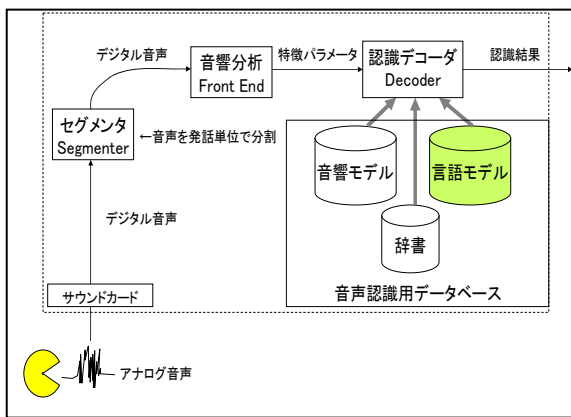


図-8 音声認識のブロック図

音声認識の仕組みを上記の図を用いて簡単に説明する。まず、アナログ音声をマイクで入力して、サウンドカードを介してコンピュータにデジタル信号として入力する。最初にセグメンタで発話されている区間を切り出して、音響分析にかけ、音声信号を特徴パラメータに圧縮する。認識デコーダでは音響モデル、言語モデル、辞書を用いて特徴パラメータの入力を文字列に変換(推定)し、音声認識結果として出力する。

以上が、音声認識ソフトウェアの機能である。既に、医療カルテ用の読み上げ音声認識ソフトウェアや議会／会議の音声認識ソフトウェアは弊社で実用化しているが、日常会話用の音声認識ソフトウェアはまだ実用化していない。これは言語モデルや音響モデルを作成するための日常会話の音声とそれを一字一句、文字に書き起したテキストデータが十分でないとい

う問題があったからである。

今回は難聴者の協力を得て、日常会話の音声を大量に集め、一字一句文字列に書き起しをして、まずは言語モデルおよび辞書を日常会話用にカスタマイズ開発した。既に会議用の話し言葉対応の言語モデルを保有しているのでその言語モデルに追加学習する方法を取った。音響モデルのカスタマイズは来年度実施する予定である。

E. 開発方法

1. 日常会話の音声収集

精度の高い音声認識を開発するためには大量の発話テキスト情報から言語モデル・辞書を作成する必要があるが、日常会話に関しては一字一句そのまま文字化したテキストデータは世の中にはほあまり存在していないという問題があった。

そこで、2010年度は日常会話や受付窓口での応対発話などの音声を幅広く収集し、一字一句書き起しをして開発のための基礎データを作成した。

日常会話の音声収集は聴覚障害者の協力を得て、その方々に対して会話する人たちの音声を収集した。

書き起しは書き起し業者(速記会社)に発注した。

以下の付録資料を参照ください。

- 付録-1 協力依頼文
- 付録-2 音声収集のお願い
(協力者への説明資料)
- 付録-3 音声収集マニュアル
- 付録-4 録音協力をお願い
(聴覚障害者と話す人への協力依頼)
- 付録-2 音声収集のお願い

2. 日常会話用の音声認識ソフトウェア

次にその基礎データを弊社保有の会議用の言語モデル・辞書に追加学習をして話し言葉用の音声認識エンジンを開発した。音声認識エンジンの開発は弊社の開発者(研究者)が行った。

言語モデルの追加学習は以下の手順で行った。

1) 言語モデル学習用テキストデータベースの作成

収集した日常会話テキストデータの文章に対し、人名をクラス(%NAME%等の記号)に置き換えた。これにより、学習用テキストに現れていない人名もユーザー辞書に追加することで、人名の文脈の時に正しく認識出来るようになる。

2) テキストデータのセグメンテーション

日本語の文章は単語で区切られていないので、形態素解析技術を用いて、単語単位に学習テキストを分割する。

3) システム辞書の作成

学習テキストに現れる単語を頻度順にソートし、基本辞書を参照して読みをふり、システム辞書を作成する。専門用語や新出単語などで基本辞書にないものは未知語として抽出されるので、必要に応じて、新規に表記と読みを人手で登録する。

4) 言語モデルの作成

セグメントされた学習テキストとシステム辞書を用いて、単語が前後にどのような単語を伴って出現するかの統計値を持った言語モデルを作成する。

5) パラメータのチューニング

実際の発話からテストセットとして抜き出したオープンな(学習データとして使用していない)音声を認識し、認識デコーダの探索幅等のパラメータを、許容できる処理速度内で認識精度が最大になるように調整する。

3. 音声認識筆談ボード(試作機)

音声筆談ボードは iPad 上に手書き入力の筆談ボードのアプリケーションを開発し、音声入

力による認識結果の文字列も表示できるようにする。この開発は弊社の開発者(システムエンジニア)が行った。

弊社では既にスマートフォン(iPhone)上で動作するアプリ、「音声認識メール」を商品化していた。これはスマートフォンに向かって音声を入力して、音声認識により文字に変換してメール文章を入力するものである。これはサーバー側で音声認識するクラウドタイプのものと、スマートフォンのみで音声認識までするスタンドアロンタイプがある。今回はクラウドタイプを利用して音声収集を行った。

iPad の発売が開始して、より広い画面表示を生かすことを考え、手書きの筆談ボードと音声認識を組み合わせることを弊社の社員が考え、この試作機開発につながった。音声認識は誤認識は避けられないので、訂正するのに、再発話やキーボード入力よりも手書き入力のほうが速くて気軽であるという発想である。

F. モニター評価

今回は音声収集、書き起しによる音声のテキストデータ化、テキストデータの追加学習による音声認識ソフトウェアの開発、および、音声筆談ボードの試作機の開発までで、モニター評価は来年度実施の予定である。

G. 開発で得られた成果

1. 日常会話の音声言語データベース

音声データ 約473時間(※1)

言語データ 上記音声データに対応する書き起したテキストデータ

※1 音声データ形式:

16Khz、16bit、speex28kbps

2. 日常会話の音声認識ソフトウェア

弊社が既に保有している、会議音声用の音声認識ソフトウェアのデータベース(辞書および言語モデル)に上記言語データを追加学習して、日常会話用の音声認識ソフトウェアを開発した。

これにより、従来から弊社が保有する会議用音声認識ソフトウェアに比べ、8ポイント(84%→92%)音声認識精度(※2)が向上した。初期カスタマイズの開発結果としては十分な成果である。また、当初目標であった、認識精度90%もクリアした。

来年度、実証実験によりモニター評価し、それにより収集する音声をデータベース化して追加学習することにより、より完成度の高い音声認識ソフトウェアにする計画である。

※2 認識精度:ここでは認識精度として、文字認識率を採用している。

文字認識率=

$$\frac{(\text{正解}\#) - (\text{誤挿入}\#) - (\text{誤削除}\#) - (\text{誤置換}\#)}{(\text{正解}\#)}$$

#:文字数

(例)

R:用紙に去年 払った医療費を書いて下さい

H:要旨見きわめ払った医療費 書いて下さい

S S S S S I D

正解文字数 (R) : 18 文字

誤挿入文字数 (I) : 1 文字

誤削除文字数 (D) : 1 文字

誤置換文字数 (S) : 5 文字

したがって、文字認識率は

$$(18 - 1 - 1 - 5) / 18 = 0.611 = 61.1\%$$

である。

3. アンケート、ヒアリング結果

今回、音声収集に協力していただいた難聴者の方々や、筆記通訳者・聴覚障害者相談員等や、厚労省に紹介していただいたアドバイザーの方々から、アンケートやヒアリングをさせていただいた。

■難聴者へのアンケート

日時:2010年10月1日
場所:横浜ラポール3階浜身連会議室
難聴者:13名

日時:2010年11月29日
場所:中野区聴覚障害者情報活動センター
難聴者:4名

アンケート内容:

(1) 日常会話支援ツールが完成したら、どのような場面で活用できると思われますか?

- 家族との会話
- 友人との会話
- 職場での上司・同僚・部下との会話
- 仕事上での取引先やお客様との会話
- 買い物の際の店員との会話
- 公共施設での受付での会話
- 病院での医師や看護師との会話
- その他

(2) 本日説明した日常会話支援ツールを使用する上で問題があるとすればどのようなことでしょうか?

(3) 以下の機器はお持ちで、日常使用されていますでしょうか?

- 携帯電話
- FAX
- iPhone
- パソコン
- インターネットメール(e-mail)

携帯端末(機種名:)

※結果は付録-5および、I. 考察3. その他の考察を参照ください。

■筆記通訳者へのヒアリング

日時:2011年2月8日(火)13:00-14:30

場所:横浜ラポール3階浜身連会議室

筆記通訳者:4名

所属:横浜市登録筆記通訳者協会

ヒアリング内容:音声認識筆談ボードに対する意見

※結果は付録-5および、I. 考察1. 試作機器(音声認識筆談ボード)を参照ください。

■聴覚障害者相談員へのヒアリング

日時:2011年2月8日(火)15:00-16:00

場所:横浜ラポール3階聴覚障害者情報提供施設

相談員:3名

所属:横浜ラポール聴覚障害者情報提供施設

ヒアリング内容:音声認識筆談ボードに対する意見

※結果は付録-5および、I. 考察1. 試作機器(音声認識筆談ボード)を参照ください。

H. 予定してできなかったこと

日常会話の音声収集を 600 時間を目指して収集したが、実績 473 時間と目標に達成しなかった。

まず、難聴者の方々に協力を依頼して、計画では 1 日 1 時間を 6 名の方をお願いすることで 100 日で収集できると見込んでいた。しかしながら、実際には最初の約 1 ヶ月 (30 日) で 15 時間しか収集できなかった。その原因は以下の通りであると分析した。

【原因】

- A. 難聴者の会話の機会が思ったほど多くない。
- B. 難聴者が話相手 (健常者) に目的を説明して録音を依頼することが困難である。
- C. 難聴者に対する発話はまだ正確に文字表示されないため会話が続き、録音のための発話が止まってしまう。

これに対し、以下の対策を行った。

【対策】

a. 健常者同士の会話も録音

これまでの収集活動から、難聴者との会話も健常者同士の会話も大差がないことがわかったため、健常者同士の日常会話音声も収集した。具体的には美容院に協力を依頼し、お客様と店員の会話を録音した。また、弊社社員同士の会話も録音した。

b. これまで弊社で収集してきた話し言葉音声も活用

模擬裁判の音声を活用した。裁判なども話し言葉の会話であるということと、裁判所でのニーズも大きかったため、有効であると判断して

活用した。

結局、以上の対策により収集できた音声は 473 時間であった。

一方、その音声を書き起す作業は約 395 時間であり、これも達成できなかった。これは上記音声収集の作業が遅れたために、書き起す作業が 3 月までに終了しなかったためである。

音声収集、書き起しの量は当初目標を達成できなかったものの、それによる認識精度の向上は当初想定 (90% 以上) を達成できたため、本事業全体の成果への影響としては問題なかったと考えている。

I. 考察

1. 試作機器(音声認識筆談ボード)

試作機の音声筆談ボードを音声収集の協力者(聴覚障害者:C.開発体制の当事者団体参照)や筆記通訳者4名、筆記通訳相談員3名等にお見せし、実際に使用していただき、意見を伺った。

その結果を【操作】、【表示】、【付加機能】に分類した。

【操作】

- ①キーボード入力よりも手書きのほうが早い
- ②指よりもタッチペンのほうが書きやすい
- ③部分的に消すモードも欲しい
- ④直前の発話を削除する機能が欲しい
- ⑤機械が苦手な人が多いので、直感に訴える簡単な操作
- ⑥気軽に利用できることが大切

【表示】

- ①字の大きさは変更できるようにしたい
- ②罫線があったら見やすい
- ③上から順番に表示できる機能も欲しい
- ④1行あけたほうが読みやすいし、手書き訂正が可能
- ⑤難聴者同士の筆談では2画面あればよい
- ⑥画面を別に大きく表示して欲しい
- ⑦アイコンをもっとわかりやすく
- ⑧次のような説明パネルがあるとよい

「私は聞こえが悪いので、ここに向かって話してください」

- ⑨表示速度をもっと早く

⑩アイコンタクトによる対話が望ましい

相手の顔を見ながら話すと理解しているかわかる

【付加機能】

- ①メモ帳などへのコピー機能がほしい
- ②外付けのハンドマイクが欲しい
- ③野外の騒音環境でも音声認識して欲しい

■考察

【操作】

- ①キーボード入力よりも手書きのほうが早い

⇒筆談ボードの有用性を確認できた。

- ②指よりもタッチペンのほうが書きやすい

⇒iPadにはタッチペンが何種類か発売されており、今回以下のタッチペンを購入して試してみたが、やはり指で直接画面上で手書きするよりもペン感覚で文字をきれいに書くことができた。特に漢字などの複雑な文字を書くときは非常に有効である。

Princeton 社製「TouchPen」



- ③部分的に消すモードも欲しい

⇒手書き色を白にして上書きすることにより、

消すことはできるが、操作が煩雑である。やはり消しゴムの機能の必要性を感じた。

④直前の発話を削除する機能が欲しい

⇒正確に発話しないと音声認識は誤認識してしまうが、その場合再度発話すると正確に変換されることがよくある。その場合、この機能は有効性である。

⑤機械が苦手な人が多いので、直感に訴える簡単な操作

⇒紙のノートに書くのと同じ感覚で操作できる今回の試作機の有効性を確認できた。ペンの太さ・色の変更やページめくり等の操作も直感的にわかるアイコン表示の工夫をしたい。

⇒【表示】⑥アイコンをもっとわかりやすく)に関係

⑥気軽に利用できることが大切

⇒直ぐに使用できるようにアプリの起動時間も早くする必要がある。

【表示】

①字の大きさは変更できるようにしたい

②罫線があったら見やすい

③上から順番に表示できる機能も欲しい

④1行あけたほうが読みやすいし、手書き訂正が可能

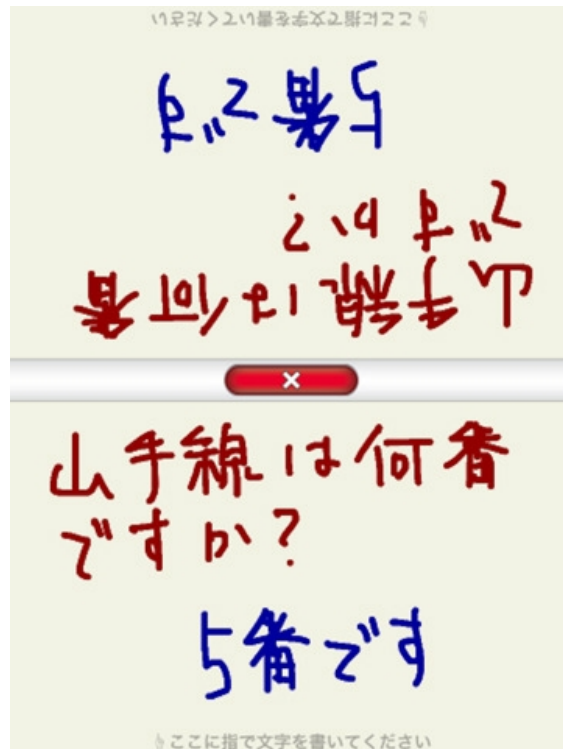
⇒これらはいずれも必要性を感じたので、採用したい。

⑤難聴者同士の筆談では2画面あればよい

⇒お互いに筆談する場合は2画面必要であろう。既に、iPad用に以下の筆談アプリが公開さ

れている。

「筆談パッド 開発者:Catalystwo Limited」



対面で筆談する際に有効であり、文字の色を変えることによってどちらが書いたものかわかりやすい。参考になる。

⑥画面を別に大きく表示して欲しい

⇒iPadの画面を外付けのディスプレイやプロジェクターに表示する機能を検討したい。

⑧次のような説明パネルがあるとよい

「私は聞こえが悪いので、ここに向かって話してください」

⇒筆談の協力を要請するために、必要な機能である。

⑨表示速度をもっと早く

⇒現状の表示速度を向上させるためには iPad 自身の処理速度向上と音声認識の処理速度を向上させることが必要である。前者は近々発売される iPad2 に期待したい。後者については音声認識サーバーの性能向上と通信回線（現状は3G もしくは無線 LAN）の速度向上が有効である。

⑩アイコンタクトによる対話が望ましい

相手の顔を見ながら話すと理解しているかわかる

⇒今回の音声認識筆談ボードではどうしても画面に目が行くので、アイコンタクトがしづらい。来年度実験予定であるメガネ装着タイプのディスプレイを検討したい。

【付加機能】

①メモ帳などへのコピー機能がほしい

⇒iPad の機能であるコピー機能を採用したい。

②外付けのハンドマイクが欲しい

⇒iPad には内蔵マイクがあるが、30cm 程度に近づいて発話する必要がある。外付けマイクの接続には以下の二通りの方法がある。

- A. iPad のヘッドフォンジャックに直接接続するタイプ
- B. iPad のドックコネクタを USB コネクタに変換して USB マイクを接続するタイプ

簡便さと音声認識精度を尺度に、最適なものを検討したい。

③野外の騒音環境でも音声認識して欲しい

⇒音声と騒音の S/N 比が小さければ（騒音が大きければ）、音声認識精度が悪くなってしまふ。やはり、外付けマイクを採用して、口元にマイクを近づける方法が確実な対策である。

2. 音声認識ソフトウェア

今回開発した日常会話用の音声認識ソフトウェアは、従来から弊社が保有する会議用音声認識ソフトウェアに比べ、8ポイント（84%→92%）音声認識精度が向上した。

認識精度を計測したテストセットは以下のとおりである。

テストセットの音声：収集した音声の書き起しテキストから 20 文を任意で抜き出したものに、日常会話で使用しそうな 16 発話のテキストを加え、男女 1 名ずつが iPad の内蔵マイクに向かって発話して録音した音声

※ただし、テストセットの書き起こしテキストは音声認識の学習テキストには含めていない。

従来の会議用音声認識ソフトの認識結果と今回開発した日常会話用の音声認識ソフトの認識結果を付録-6に示す。

従来の会議用音声認識ソフトでは誤認識していたが、今回開発した日常会話用の音声認識ソフトでは正しく認識した例を以下に示す。

[凡例]

（正解）実際に発話した文章

（従来）従来の会議用音声認識ソフトで認識した結果

（今回）今回開発した日常会話用音声認識ソフトで認識した結果

■例 1

(正解)今晚の晩ご飯は何を食べたいですか

(従来)今般の単語版は何を食べたいですか

(今回)今晚の晩ご飯は何を食べたいですか

「晩ご飯」という単語は会議用音声認識の学習テキストにはほとんど含まれていなかったが、今回収集した音声は日常会話であったため、この単語も含まれて学習されたため、正確に認識されるようになった例である。

■例 2

(正解)すき焼き用のお肉はデパートの地下で買ってきてもいいですか

(従来)月昭夫 のに加えて窓口課で できてもいいですか

(今回)すき焼き用の 肉はデパートの地下で買ってきてもいいですか

「すき焼き」、「デパート」という単語も同様に会議用音声認識の学習テキストにはほとんど含まれていなかったが、今回収集した音声は日常会話であったため、この単語も含まれて学習されたため、正確に認識されるようになった例である。また、(従来)は「デパート」が認識されなかったため、それに続く「地下」も誤認識しているが、(今回)は正しく認識している。

■その他の改善例

(正解) (従来) (今回)

不安 → ふあふあ → 不安

お腹 → 中 → お腹

葛根湯 → 買った本当 → 葛根湯

酒に酔ってないとき → 先にやって内的 → 酒に酔ってないとき

.....

いずれも、会議では話さないが日常会話では話す単語がほとんどである。

3. その他の考察

■難聴者へのアンケート に関して結果を考察する。

図-9は「1. 日常会話支援ツールが完成したら、どのような場面で活用できると思われますか?」の集計結果である。

やはり、「医師や看護師との会話」が一番多く、全員が必要と答えている。弊社では医療の電子カルテ入力用の音声認識ソフトウェアも開発しているため、それを応用して開発することは可能であるが、モニタテストをどのように実現するかが課題となる。関係団体と協議したいと思う。

「友人との会話」、「家族との会話」、「受付(病院や役所)での会話」、「買い物での会話」がそれに続く。「友人、家族」の二つは日常生活で最も会話の機会が多いからであり、「受付」は機会の多さよりも必要性が高い場面であると考えられる。これらの会話は来期、重点的にモニタ実験を行いたいと思う。

一方、職場や取引先との会話については少なかった。これは今回の協力者が職場を退職されている方が多かったのが原因と考える。来期は職場でのモニタ実験ができる協力者にも依頼したいと思う。

また、「その他」では「習い事、趣味のサークル、

町内会等の会合、子供の学校での教師や父兄、TV、ラジオ内容の理解」などがあつた。今回の開発対象は 2,3 名での日常会話をターゲットにしているので、会合など、複数の人々が発言する場面では、特に、健常者の人々が主体の場合は、マイクの設置や誤変換したときの対処方法など、困難な課題を克服する必要がある。これに関しては、弊社で開発している会議議事録作成支援システムの活用などの方法で別途、検討したいと考えている。

「TV、ラジオ内容の理解」は切実な要望である。今回のような東日本大震災の際の緊急放送の可視化は非常に重要な課題である。これは放送局等とのタイアップが必要である。関係団体と協議したい。

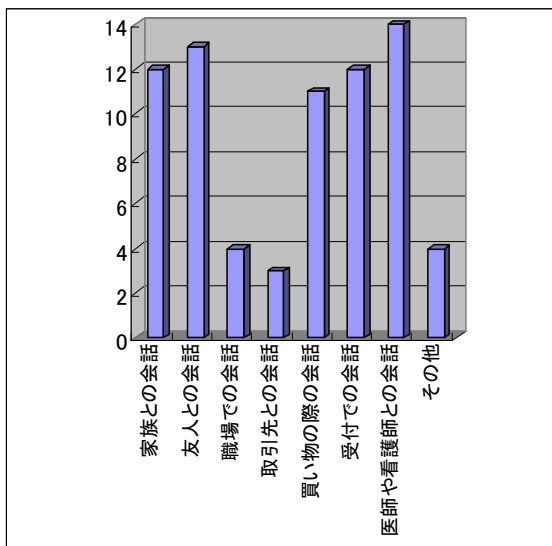


図-9 アンケート集計結果(1)「1. 日常会話支援ツールが完成したら、どのような場面で活用できると思われますか?」

■「4. 日常会話支援ツールを使用する上で問題があるとすればどのようなことでしょうか?」については以下のような回答があつた。

①初対面の人にマイクを向けるのは慣れるまで抵抗がありそうです。

②相手の都合、許可

③話者の許可

④認識精度が一番問題だと思います。会話の時間がクリアできるか心配です。

⑤会議などでマイクを持っていない人の発言を拾うことが難しい。マイクを真ん中に置いて、多方向から会話を拾うことができれば嬉しいと思う。

⑥騒音中での会話変換

⑦マイクの使用

⑧私は高齢で一人暮らし。会話が少ないのと機械の取り扱いがスムーズにできないかも

⑨費用の点

①②③は、話し相手に音声入力の同意を得たり、その必要性を説明して納得していただくことが難しいもしくは労力がかかると感じているということだと考えられる。相手に意識せず気軽に使用していただく工夫が必要である。

④⑤⑥は音声認識精度の問題である。④は音声認識精度を高くして、修正しなくても意味が通じれば会話がテンポよく進むが、誤認識のために意味不明であればその修正で会話が中断してしまう問題が大きい。⑤は指向性を制御できるマイクロホンアレイの技術が有効かもしれない。⑥はノイズリダクション技術や咽頭マイクなどの技術の応用が考えられるが、まだ実用化されていない。今後の課題である。現状では⑦のマイクを使用し、口元の距離を10cm以下にして、クリアな音声を入力して認識精度を向上させる方法が現実的である。

⑧は次回のモニタ試験で使いやすさを追求し

たい。

⑨の費用は買取ではなく、月額サービスで気軽に利用できる料金体系を考えたい。

■「5. 以下の機器はお持ちで、日常使用されていますでしょうか?」については、図-10のような結果であった。

FAX はほぼ全員が保有しているが、携帯電話、パソコン、インターネットメールの保有率も非常に高い。

一方、携帯端末を持っている人は 2 名のみであった。これは一般の人も同様かもしれないが、今回の試作機の音声認識筆談ボードのようなアプリケーションが日常生活で活用できるのであれば、保有率も高くなるのではないかと思う。

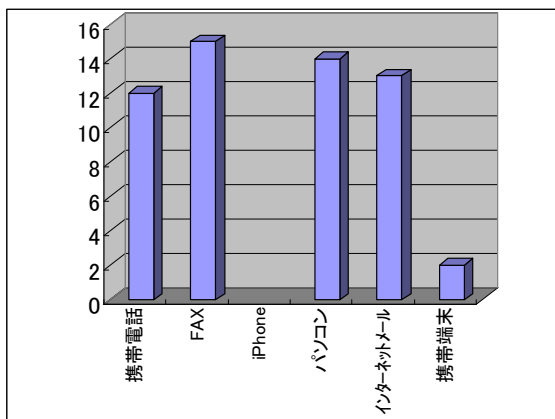


図-10 アンケート集計結果(2)「5. 以下の機器はお持ちで、日常使用されていますでしょうか?」

J. 結論

聴覚障害者の日常生活支援機器として、音声認識し文字表示する携帯可能な機器の開発を行った。本年度は日常会話用の音声認識ソフトウェアを開発するための基礎データとして音声を収集し、一字一句の書き起しを行った。また、その基礎データを使用して日常会話用の音声認識ソフトウェアを開発した。さらに試作機として手書き入力もできる音声認識ボードのプロトタイプを iPad 上で開発した。

音声収集に関しては目標時間を 600 時間と設定したが、最初、聴覚障害の人たちとの会話と限定したために、思ったように集まらなかったが、健常者同士の日常会話と大きく変わらないということで、音声収集の対象を広げ、473 時間収集することができ、書き起しも 395 時間分の音声を文字化することができた。

上記音声収集の時間は目標を下回ったものの、それを基礎データとして開発した言語モデルを使用した音声認識ソフトウェアの音声認識精度は 92%を達成し、目標の 90%を超えることができた。来年度は試作機のモニタ使用をしながら、実際の聴覚障害者との日常会話の音声収集を継続し、学習データを増強し、書き起されたテキストデータは言語モデルの追加学習データとし、更に音響モデルの学習にも使用して、認識精度の向上を図る。

試作機の音声認識筆談ボードに関しては、聴覚障害の方々や、筆記通訳や相談員などの関係者からのヒアリングにより、貴重な改善意見をいただいた。その結果を踏まえて、来年度、試作機を改良して、モニタ使用(実証実験)と改良の繰り返しを行って、実用機として完成させる予定である。

K. 健康危険情報

1. 開発者側
特に無し。
2. 当事者側
特に無し。

L. 成果に関する公表

1. ホームページ、刊行物等の紙面などでの発表

特に無し。

2. 展示会などでの発表

平成 21 年度障害者自立支援機器等研究開発プロジェクト研究成果一般公開で発表し、NHK の取材を受け、2011 年 3 月 9 日の 17 時のニュースで放映された。

M. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得
特に無し。
2. 実用新案登録
特に無し。
3. その他
特に無し。

以上

開発成果の公表に関する一覧表

(展示会など)

発表者氏名	展示会名	主催者	開催期間	開催場所
中村雅巳	平成21年度障害者自立支援機器等研究開発プロジェクト研究成果一般公開	厚生労働省	平成23年3月9日	厚生労働省

開発成果の公表に関する刊行物・別刷

特になし。

付録－ 1
協力依頼文

社団法人全日本難聴者・中途失聴者団体連合会

理事長 高岡 正 殿

2010年9月16日

株式会社アドバンスト・メディア

代表取締役社長 鈴木清幸

音声データ収集および実証実験の協力依頼

拝啓

貴会、ますます盛栄のこととお喜び申し上げます。

さて、弊社は「平成22年度障害者自立支援機器等開発促進事業 聴覚障害者の日常生活支援機器の開発」事業に応募し、採択内示を受けました。開発テーマは「音声認識し文字表示するメガネや携帯可能な支援機器の開発」です。

今回の開発目的は、聴覚障害者の方が日常生活で遭遇するさまざまな場面での会話によるコミュニケーションを支援するために、対話者の発話を音声認識して文字表示する携帯端末を開発することにあります。そのためには、実際に聴覚障害の方々が必要とする対話の音声データを収集し、音声認識エンジンを開発することが大切です。

そこで、貴会に、音声データの収集や実証実験の協力をお願いしたいと考えております。

別紙に協力依頼内容の概要を記載いたしましたのでご確認ください。

ご多忙とは存じますが、ご協力のほど、よろしく願いいたします。

敬具

聴覚障害者支援機器開発に伴う音声データ収集および実証実験の協力依頼内容

2010年9月16日

株式会社アドバンスト・メディア

開発代表者 中村雅巳

【本開発事業の目的】

日常生活において、人と人のリアルタイムなコミュニケーションは音声による会話によって行われることが多いが、聴覚障害者にとってはその音声言語がコミュニケーションのバリアとなっている。その不便を解消する支援機器を開発するのが本事業の目的である。

【事業の実施概要】

■開発する支援機器

携帯電話 iPhone、携帯コンピュータ iPad 上で動作する音声認識文字表示アプリケーション

■想定している使用方法

聴覚障害の人が持ち歩き、日常会話の際に対話相手の健常者が iPhone に向かって発話して文字表示し、それを聴覚障害の方が読んで理解する。

■開発スケジュール

2010年度：音声認識技術開発に必要な実際の日常会話の音声収集、テキスト書き起し

2011年度：支援機器の開発

実証実験

【依頼内容】

●iPhone（音声認識機能付）を貸与いたしますので、持ち歩いていただき、日常生活の中で対話相手に iPhone を渡し、それに向かって発話していただき、その音声を録音してください。

録音の際には、まだ未完成で誤変換がありますが音声認識して文字表示されますので、多少なりとも会話支援に役立てていただきたいと思います。

●音声収集は複数の人が発話する会議場面ではなく、一対一での対話場面で収集してください。

●対話相手の発話音声のうち、「あなた」に話している発話をすべて録音願います。

ただし、知られたくないプライベートな情報は対象外としていただいで結構です。

●音声収集者の居住場所が標準語使用地域で、標準語を収集してください。

収集に当っては、別紙、収集先の場面表を参考に幅広く収集してください。

●1日あたり、録音音声は1時間以上を目標に収集していただきます。

(1)水曜日にお渡ししますので、木曜日から月曜日まで5日間、音声を収集していただく。

(2)火曜日にiPhoneを返却していただき、その週の木曜日からは別の人に音声収集を依頼します。

(3)翌週の水曜日に再びiPhoneをお渡ししますので、木曜日から月曜日まで5日間、音声を収集していただく。

※1日1時間、5日で5時間以上収集できなかった場合も火曜日に一旦、返却していただきます。

●すなわち隔週、音声を収集していただき、録音音声の合計時間が目標時間になれば終了します。目標時間は現在のところ、50時間を見込んでいますが、状況により変更する可能性もあります。

●協力していただける方12名を二つのグループに分けます。

Aグループ：6名

Bグループ：6名

したがって、各グループの方々には以下の順番で音声を収集していただきます。

(1)木曜日～月曜日：Aグループの音声収集

(2)火曜日にiPhoneを返却していただき、音声をサーバーに回収

(3)水曜日にiPhoneをBグループに送付

(4)木曜日～月曜日：Bグループの音声収集

(5)火曜日にiPhoneを返却していただき、音声をサーバーに回収

(6)水曜日にiPhoneをAグループに送付

(1)～(6)を10回ほど繰り返すことになります。

以上の方法は実施していくうえで改善すべき点が発生すれば、適宜変更していきたいと考えています。

【協力者の条件】

1 社会に出て、積極的に健常者と交流している人(例えば、難聴協会役員として地方公共団体

や他団体の健常者と色々交渉などしている人)で、健常者の音声を集めることが出来る人。
又は企業等で、仕事をしている人で、社内での日常会話の音声を集められる人。
(社内の秘密を集める必要は無い。)

2 自分のPCを所有してメールが出来る人。

弊社からメールで各種連絡いたしますので、メールは必須となります。

ファイル等の送付の可能性もございますので、インターネットに接続できるパソコンでのメールが必要です。必要に応じてメーリングリストに登録していただきます。

3 携帯電話を使用している人、iPhoneを使用できる人が望ましい。

【その他応募に関する事項】

1 状況説明、情報共有のための集合・会議等は横浜ラポール会議室で実施します。

2 事務的なことでアドバンスト・メディア社に集まることもあります。

(東京都豊島区東池袋 3-1-4 サンシャインシティ文化会館 6階内)

3 収集の場面或いは収集の方法の見直しをする可能性もあります。(次頁作業表参照)

4 本作業の謝礼は多くはありませんが、収集した音声収録時間に見合ったアルバイト程度の経費は支払われます。

【個人情報の機密保持について】

弊社(株)アドバンスト・メディアはプライバシーマークを取得しているため、録音した音声や書き起しをしたテキストデータの利用に関して、特に個人情報の取り扱いは万全の体制で臨んでいます。そのことに関して記載した別紙の同意書の内容をご確認の上、署名をいただきます。

以上

付録－２
音声収集のお願い
(協力者への依頼説明資料)

日常会話収集のお願い



- 難聴者のための日常会話支援として、音声を文字変換して表示する携帯機器を開発することが目的です。
- 正確に文字変換するためには、たくさんの日常会話の音声を収集して音声認識エンジンに学習させることが必要です。
- 今回は皆様方に協力していただき、日常生活でのさまざまな会話の音声を録音・収集していただきたいと思っております。

1

誰の会話を録音するか

- あなたに話しかける人の音声を録音してください。
 - 1. 家族
 - 2. 友人
 - 3. 職場の同僚
 - 4. 店員(買い物や飲食)
 - 5. その他
- あなたに話しかける人すべてが対象になります
- ただし、プライバシー的に不都合な内容であれば録音はしなくて結構です。



2