

音声認識を利用したスケジューラアプリの予備実験

服部哲^{†1} 柴田邦臣^{†2}

概要: 話し言葉によるコミュニケーションやこれからやるべきことを覚えておくことが困難な人や子どものために、筆者らは音声認識を利用してコミュニケーションを支援したりスケジュール管理を支援したりするタブレット端末用アプリを開発している。これまでに本アプリの主要機能を開発してきたわけであるが、本論文では試作アプリを利用して行った予備実験について述べる。本予備実験では仙台市の NPO 法人に協力していただき、本アプリの使用感や有用性を評価していただくとともに、改善すべき問題点を明らかにすることを目的とした。予備実験の結果、本アプリの有効性を示唆する意見が得られた。また、追加・拡張すべき機能も明らかになった。

キーワード: スケジューラ, コミュニケーション支援, タブレット・アプリ, 音声認識, ピクト

Preliminary Evaluation of a Scheduler Application using Voice Recognition

AKIRA HATTORI^{†1} KUNIOMI SHIBATA^{†2}

Abstract: We have been developing a tablet application for the children and persons with deaf or hard-of hearing and a developmental disability. Such children have difficulties communicating using spoken language and keeping what will happen next in mind. Our tablet application enables them to communicate with hearing people and to understanding a sequence of activities that will occur. The application shows a sequence of the pictures and words corresponding to conversation using voice recognition. We conducted a preliminary experiment of the application in cooperation with a NPO. In this paper, we present a brief description of the application and describe the preliminary experiment.

Keywords: Scheduler, Communication support, Table application, Voice recognition, Pictures

1. はじめに

ほとんどの人は話し言葉を使ってコミュニケーションを取っている。社会はコミュニケーションから成立すると言う考え方からも、言語を獲得し、話し言葉でコミュニケーションすることの重要性・必要性を理解できる。しかし、難聴や知的障害、自閉症などの障害によって話し言葉を自由に使えない人たちがいることも事実である。そのため、視覚シンボル（本研究ではピクト）を利用したコミュニケーション支援の研究が行われている[1]。視覚シンボルを利用したコミュニケーション支援は AAC (Augmentative and Alternative Communication: 拡大・代替コミュニケーション) の研究領域であり、今日では AAC を実現するスマートフォンやタブレット向けのアプリも開発されている（文献[2][3][4]を参照）。

一方、活動の予定をカレンダーや手帳で管理することは一般的であるが、活動の見通しを持つことが困難な人がいることもまた事実である。そのため 1 日や 1 週間の予定をスケジュールボードに掲示することによって活動の見通しを持てるようにすることが行われている[5]。

筆者らもこれまでに難聴児・者の言語獲得やコミュニケ

ーションを支援するためのタブレット・アプリの研究開発に取り組んできた[6][7][8]。本アプリは音声認識によって聴者の話し言葉を取得し、それを写真やイラスト（本研究ではピクトと呼ぶ）と文字を組み合わせ提示する。ピクトによってスケジュールを示すことで、難聴児・者のコミュニケーションを支援し、活動の見通しを持つことが困難な人の状況理解を促進することを目指している。これまでにアプリの基本機能を開発してきたが、評価実験が不十分である。そこで本論文では NPO 法人の協力で行った予備実験について述べる。音声認識技術を利用して聴覚障害者・児のコミュニケーションを支援する技術の開発が期待されており、本研究はその期待に貢献するものである。

2. 関連研究

聴覚障害者を対象とした情報保障技術の一種として、音声認識を利用した字幕作成の自動化についての研究が行われている[9][10]。筆者らの研究も字幕についての研究から示唆を受けている[6]。

文献[11]では広汎性発達障害児・者の会話理解を支援するコミュニケーションエイドについて述べている。このシステムは音声認識した会話相手の発言を視覚シンボルに変換し、それを 5W1H の観点で構造化するものであり、モバイル端末上のアプリケーションとして開発されている。当事者とその支援者による短期と長期の利用実験を行い、提

^{†1} 駒澤大学
Komazawa University
^{†2} 津田塾大学
Tsuda College

案アプリケーションの有効性を示した。文献[12]では聴覚障害者の日常会話を支援するために、聴覚障害者と対話する相手の発話を音声認識し、文字で提示する機器の開発している。磁気で何度も手書き可能な筆談ボードをタブレット上で動作するアプリケーションとして実現したものである。このアプリは音声認識が誤認識しても手書きで修正可能であり、タブレット端末上で動作する。システム開発と同時に日常会話の音声収集も進め、音声認識ソフトウェアの精度向上も行っている。音声認識はインターネット上のサーバが行う。現在「UD 手書き」としてアプリが公開されている。

聴覚障害者支援のためのマルチモーダル発話可視化に関する研究も行われている[13]。この研究では会議のような複数人かつリアルタイムな対面コミュニケーションを想定し、聴覚障害者が聴者の発話内容を理解できるようにするため、発話内容を音声認識処理を通してテキスト化して提示したり、発話者の口元の映像を表示したりするなど、複数の情報の提示・可視化方法を組み合わせた Web アプリケーションを開発している。

一方、文献[14]では時間管理の概念を持っていないことが多い自閉症児のためのスケジューラアプリを研究開発している。このアプリでは利用者である自閉症児に達成感を感じる機会や彼・彼女自身で何かをやってみようとする動機を与えることを目指している。そのため単純な To-Do 管理だけでなく、何かを達成したときにスタンプを与えたり、To-Do の手順を詳細に示したりする機能を備えている。特別支援学校と協力し、このアプリの有効性を明らかにした。

音声認識は利用されていないが、視覚シンボル（日本語 PIC）を利用することで、日本語の文字を使わずにメールをやり取りすること可能にするシステムも研究されている[15]。このシステムは Web アプリケーションとして開発されており、利用者はシンボルを並べることによってメッセージを作成することができる。

これらに対し筆者らのアプリは音声認識を全面的に利用し、話し言葉をピクトの並びでスケジュールを提示することによって、難聴児・者と聴者とのコミュニケーションや活動の見通しを持てるように支援する。

3. アプリの概要

筆者らはこれまでに難聴児・者を対象として言語獲得やコミュニケーションを支援するためのタブレット端末向けアプリを研究開発してきており[x][x]、活動の見通しを持つことが困難な人や子どものためにスケジューラアプリへと発展させてきた[x]。現段階では Android タブレットを対象にアプリ開発を進めている。

本アプリは難聴児・者や発達障害の人たちの支援者（聴者）の発話内容を音声認識し、その結果として取得できた文字列に対応する絵や写真（ピクト）を検索し、タブレッ

ト端末の画面上に次々と表示する。

一連のピクトをシーケンスと呼んでおり、「目標設定」と発することによってシーケンスを予定や To-Do として登録し、スケジューラとしての活用することも検討してきた。

本アプリの機能は(1)ピクト登録、(2)シーケンス作成、(3)シーケンス閲覧・編集からなる。ピクト登録ではタブレット端末のカメラを利用し、その場で撮影した写真にタグを付与してピクトとして利用できるようにする。シーケンス作成では音声認識の結果として取得された文字列によってピクトを検索する。音声認識の結果はテキストビューに保持されるため、そのテキストビューを決められた方向へスワイプすることによってもピクト検索を実行することができる。作成されたシーケンスは音声認識の結果とともにタブレット端末に保存される。シーケンス閲覧・編集では保存されたシーケンスを表示し、ピクトの並びを変更することができる。

図 1 は本アプリのユーザインタフェースである。図中の①はコンソールであり、各機能呼び出すことができる。同②はシーケンスである。本アプリのインストール時に登録されるピクトや利用者によって登録されたピクトが表示される。同③には音声認識の結果として取得された文字列が表示される。音声認識の結果は形態素解析され、名詞や動詞、形容詞だけがテキストビューに表示される。テキストビューをタップすることによってコピーが作成され（図 1 中央下の緑色のテキストビュー）、自由にスワイプできる。このコピーされたテキストビューを操作することによって新たにピクトを登録したりすることができる。

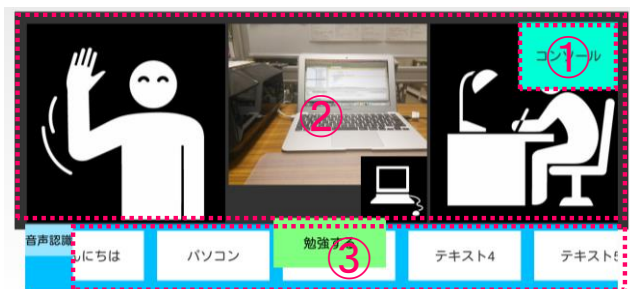


図 1 本アプリのユーザインタフェース

Figure 1 User interface of our tablet application.

これまでアプリの基本機能は実装された。しかし評価を行っていない。本論文では予備実験について述べる。

4. 予備実験

4.1 目的と方法

筆者らは本アプリを一般公開し、国内外の福祉分野に寄与することを目指している。そのため本アプリを公開できるレベルまで高めなければならない。そこで本研究では、本アプリを実際に利用していただき使用感や有用性を確認

し、その中で改善必要点や問題点を明らかにするために、まずは少人数の協力者による予備的な評価実験を行った。

本研究では NPO 法人ビートスイッチ（所在地は仙台市）と協力し実験を行った。具体的には、発達障害や難聴のため言葉でのコミュニケーションが困難な人の支援者 4 名（以下、協力者 A, B, C, D とする）に本アプリをインストールしたタブレット端末を貸し出し、自由に利用していただいた。貸し出した端末のホーム画面には本アプリのアイコンだけを表示するようにした。いずれの端末も Wi-Fi モデル。協力者 D は自宅にインターネット環境がなかったため、モバイル Wi-Fi ルータも貸し出した。

協力者 4 名のうち 3 名が発達障害、1 名が難聴・弱視である。実験期間は 2015 年 7 月 22 日から 9 月 7 日である。実験に先立って本アプリの使い方を説明し、協力者全員がアプリの使い方を一通り学習した。マニュアルも配布した。

協力者には従来型の携帯電話（いわゆるガラケー）を利用している人もいれば、スマートフォンを利用している人もいたが、いずれにしても電話とメールが利用目的であった。自宅でパソコンを使う人もいれば、NPO 法人の事務所にいるときだけパソコンを使う人もいた。

タブレット端末の貸し出しと同時にアンケート用紙も配布し、実験期間終了後に回収すると伝えた。アンケートでは使い勝手や有用性など本予備実験の目的を踏まえて表 1 に示した質問をした。質問 1 と質問 2 および質問 4 では 5 項目の選択肢から当てはまるものを選択していただき、また自由回答も得るようにした。このアンケート用紙による調査に加え、対面で聞き取り調査も行った。

表 1 アンケート内容
Table 1 Questionnaires.

1. 使い勝手は、いかがでしたか？ 感想をどうぞお書きください。
2. このアプリは、どのような場面で役立ちそうですか？ 生活の中で使えそうな場面がありましたら、お教えてください。
3. このアプリで不足だったところ、必要な改善点を、教えてください。
4. 改善が果たされた場合、このようなタブレット・アプリを、どのように使いたいですか？ こんな使い方ができそうだと、という点を教えてください。
5. そのほか、気がついたことがありましたら、自由にお書きください。

4.2 ピクトと音声認識方法

本予備実験では、筆者らが日常で利用する中で必要であろうと考えた 444 個のピクトをあらかじめアプリに登録した。あらかじめ付与したタグの数は 4455 個である。

音声認識の方法として、Android に付属の音声認識機能を利用する方法や外部サービスを API 経由で利用する方法などがあるが、本研究では常時音声認識しているよりも、音声認識したいときに利用者が指定するほうがわかりやすいであろうと考え、Android に付属の音声認識ダイアログを呼び出すようにした。

4.3 結果

協力者 A と B は自宅にインターネット接続環境があったがタブレット端末をインターネットに接続することができなかった。協力者 A の代わりに新たな協力者に端末を貸し出し、現在も利用を継続していただいている。

協力者 C は実験期間中に家庭の事情でアプリの利用が困難になったため、タブレット端末を回収した。協力者 C が本アプリを利用可能になり次第、利用を再開していただく予定である。

協力者 D はモバイル Wi-Fi ルータも貸し出したためアプリの利用が可能であったが、アンケート用紙は回収できていない。協力者 D のタブレット端末は回収できたものの、保存されたシーケンスはゼロであった。

予備実験ではアンケート用紙の回収はできていないが、実験期間終了時やタブレット端末の回収時に聞き取り調査によって意見を得た。それらの意見を整理する。

まず本アプリの効果や有用性を示唆するものとして、

- 身の回りの話は、しやすくなったところがある
- 「一緒に○○（アプリのこと）しよう」というのが言いやすい
- 「これから○○（予定のこと）する」というときも使いやすい
- ピクトの組み合わせがおもしろい

これらの意見があった。「ピクトの組み合わせがおもしろい」というのは、ピクトに付与されたタグの意外性についての意見である。一方、改善点を指摘する意見は以下のようものが得られた。

- 用途として、移動中に使いたいが、たいていネットが切れてしまう。イーモバでもだめ
- 「お父さん」「お仕事」など重要な用語が出ない
- 「ちょっと外で使う」というのができない。家は家で、もう方法があるので
- 指文字を使ったりするので、ピクトがない分は、手話の画像をいれておいたらどうか？
- 手話なら、動画でいれてしまえとよいが、どうか？
- 画像を出すとき、と、画像を登録するとき、がわからない
- すぐ混乱してしまうので、使いやすいデザインにしてほしい

- 時計のピクトをいれてしまったらどうか？
- それぞれのイラストを入れておいても良いかも
- 事前に写真やイラストをいれておく、というのはどうか？
- 立てた予定ができたなら、花丸がもらえるようにするとよい
- タイマーが着けられると良い。それは砂時計だとさらによい
- 「課題設定」「課題達成」と言うのと「なままる」になるとか
- 予定を入れて、その予定に誘い出せるように

あらかじめ登録するピクトについて指摘や、楽しく予定が立てられ、それが後から確認できるようにするなど、示唆に富む意見が得られた。ピクトは事前に登録されているものの、「お父さん」「お仕事」など重要な用語が出ない理由は、音声認識がうまく機能していないためと思われる。

5. 考察

実験協力者からの意見では、身の回りの話をしたり、これからの予定を示したりするときに、本アプリが有効活用されることが示唆された。つまり難聴や発達障害などのため話し言葉によるコミュニケーションが困難な人たちや、これからの予定に見通しを持つが困難な人たちに役立つ。たとえば集中しやすくなるなどのメリットがある。今後さらにアプリの完成度を高め、一般公開するために必要な改善点について以下で議論する。

(1) 時間管理と課題設定

本アプリでは「課題設定」と発することでシーケンスを To-Do として登録することができる。しかしその To-Do を達成できたのかどうか、また達成したとしてもアプリを通して利用者が達成感を感じられる仕組みがない。また、To-Do の時間を管理する機能やタイマー機能もない。実験協力者が指摘しているように、時計の画面をピクトとして登録したり、砂時計のようなタイマー機能を追加したりするなど。「今日何時にどこで何をやる」「明日の何時にどこで何をやる」などの予定や課題を登録し、これらを実行できた場合はアプリの画面上で目立つように「はなまる」などのマークが付与される。これによって楽しく予定を立てることもなる。

(2) オフラインの音声認識

本研究では Android に備わっている音声認識機能や API として提供されている音声認識機能を利用した。これらはいずれもインターネット接続を前提としており、オフラインで利用することができない。実験協力者のコメントにあるように、普段自宅ではコミュニケーションの方法やルールが確立されているため、本アプリを使うことはほとんど

ない。実際は移動中や移動先での利用が多くなると予想される。そのためには、インターネット接続がなくても本アプリを利用できるように、オフラインの音声認識を導入することが必要である。また、タブレット端末をインターネットに接続することができない利用者もあるため、その課題に対してもオフライン音声認識は効果的である。Julius などの音声認識エンジンを組み込むことや AmiVoice などの音声認識専用の開発キットを導入するなどの解決策がある。しかし、導入の技術的ハードルや金銭的成本を考えるとこれらの解決策はどれも決定打となりにくい。したがって本研究ではオフライン音声認識よりも他の改善点を優先的に開発していく予定である。

(3) ピクトの管理

本アプリではインストール時に登録されるピクトに必要なものが存在しない場合、利用者がタブレット端末のカメラを利用して撮影した写真をピクトとして登録することができる。あらゆる利用者が日常的に利用するであろうピクトをすべてあらかじめ登録しておくことは不可能であるし、家族の顔写真などは個々の利用者ごとにカスタマイズが必要である。そのため、このピクト登録機能は本アプリに不可欠な機能であるし、実験協力者もこの機能を利用してピクトを登録している。しかし、難聴であり、かつ弱視でもある実験協力者からピクトを拡大表示する機能が必要であることが指摘された。

この課題について筆者らは機能拡張を進めている。具体的には、Android の ImageView を拡張し、2本の指を利用したピンチインやピンチアウトによる拡大縮小可能なビューを新たに開発した。本アプリではシーケンス表示領域（図1の②）に表示されたピクトをダブルクリックすることによって、そのピクトを拡大縮小可能なビューにコピーするようにした（図2）。

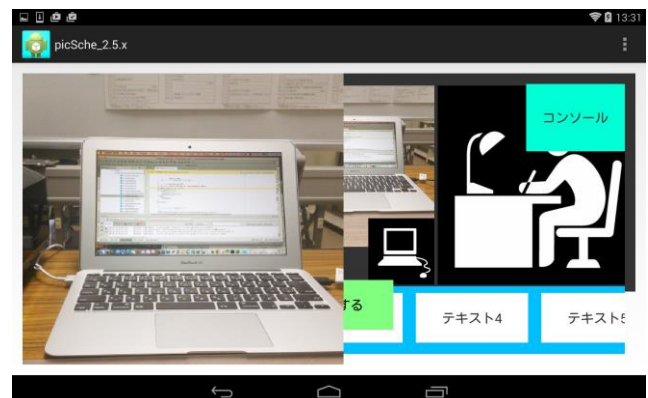


図2 ピクトの拡大表示

Figure 2 Zoom display of a picture.

ところで、言葉によるコミュニケーションが困難な人たちはすでに視覚シンボルを利用してコミュニケーションし

ていることもある。したがって、それらの視覚シンボルを本アプリに簡単に取り込むことができれば効果的であると思われる。実際、実験協力者からもその必要性が指摘された。そのため、たとえば手持ちの視覚シンボルをスキャンしたり撮影したりするなどしてデータ化し、その画像データをサーバにアップロードし、簡単にタグ付けすることができ、さらにそのタグ付けされた画像データを本アプリにダウンロードして利用可能にするような機能やサービスの開発を検討している。

(4) ユーザインタフェース

今回の実験では本アプリが積極的に活用されたとはいえない。その大きな原因はインターネット接続できなかったことであるが、もう1つ、シーケンス作成やピクト登録の機能の画面までたどり着けなかったことがあげられる。具体的には、本アプリを起動したときに表示される画面にはコンソールボタン(図3の○)と音声認識の結果として取得された文字列が表示される領域だけが表示されており、利用者がコンソールをタップしてシーケンス作成やピクト登録を選ぶことで、初めてそれらの機能が利用可能となる。そのため、本アプリの使い方を説明したとはいえ、タブレット端末の操作や本アプリに不慣れな利用者にとってこれらの機能に画面までたどり着くことができなかったと思われる。そのため本研究ではアプリの画面遷移を再設計し、シーケンス作成とピクト登録を本アプリの2大機能と考え、それらの機能を中心とした画面遷移、つまりアクティビティの切り替えを実装している。そのためこれまでの実装方法を見直し、すべての機能のアクティビティ化を進めている。

一方、本アプリでは利用者が登録したピクトをシーケンスに表示する際、インストール時に登録されたピクトの中に同じタグのピクトが存在すれば、後者を利用者が登録したピクトの右下に小さく重ね合わせて表示するようにしている(図3)。これは言葉によるコミュニケーションが困難な子どもたちの概念理解を支援することにも役立つと考えている。図3のようにピクトが重ね合わせられた場合は問題ではないが、右下に小さく表示するピクトが利用者によって登録されたピクト(写真)に映っている大切な箇所を隠してしまうことがあると指摘された。ピクトの拡大縮小機能の導入は1つの解決策であるが、毎回拡大縮小するのは煩わしいため、ピクト表示方法については再検討の必要がある。

6. おわりに

筆者らは難聴や発達障害あるいは高齢のために、話し言葉によるコミュニケーションが困難であったり、相手の意図を理解することが困難であったり、自分のやるべき行動を覚えておくことが困難であったりする人や子どものため、音声認識を利用したコミュニケーション支援やスケジュー

ラアプリの開発を進めている。そして本論文では、仙台市のNPO法人に協力していただいたスケジューラアプリの予備実験について述べた。本予備実験の結果、本アプリが難聴や発達障害の人や子どもにも効果があることを示唆する意見が得られた。しかしながら、今回の予備実験では協力者は十分にアプリを利用することができなかった。そのため、今後は今回の実験で指摘された問題点を解決し、より多くの協力者が得られる段階までアプリの完成度を高める必要がある。その上で、本アプリの評価実験を行い、有効性を明らかにしなければならない。これらの研究課題をクリアし、本アプリを一般に公開することが本研究の目標である。

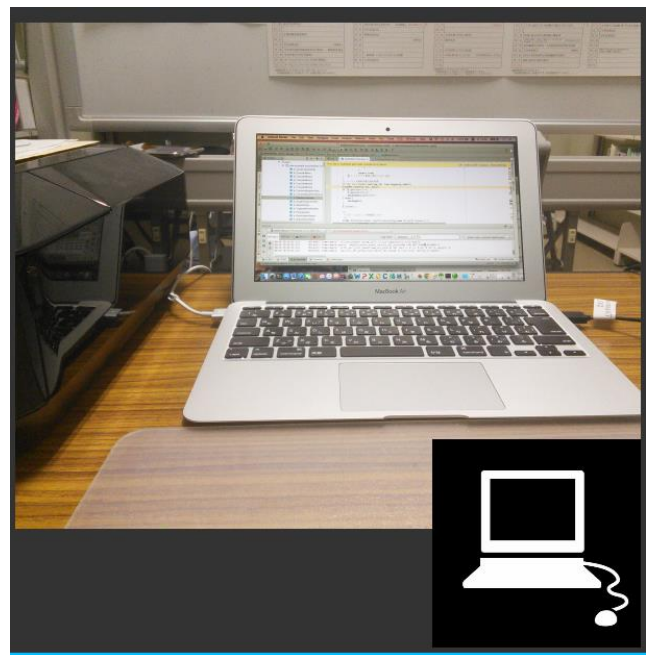


図3 ピクトの重ね合わせ表示

Figure 3 Overlaying of pictures.

参考文献

- 1) 藤澤和子, 桜井貴文: 視覚シンボルによるコミュニケーション支援に関する研究, 風間書房 (2011).
- 2) 岡耕平: 障害支援技術としての携帯情報端末アプリの分類の試み—発達障害のある人への支援を中心として—, 電子情報通信学会技術研究報告, Vol.113, No.72, pp.135-140 (2013).
- 3) 東京都障害者IT地域支援センター: iPhone, iPad用・障害のある人に便利なアプリ一覧, <http://www.tokyo-itcenter.com/700link/sm-iphon4.html>
- 4) 東京都障害者IT地域支援センター: Android携帯用・障害のある人に便利なアプリ一覧, <http://www.tokyo-itcenter.com/700link/sm-and1.html>
- 5) 金森克浩(編著): 特別支援学校におけるATを活用したコミュニケーション支援, ジアース教育新社 (2010).
- 6) 柴田邦臣, 阿由葉大生, 服部哲: 言語獲得と知識獲得を支援するタブレット・メディア—聴覚障害児のコミュニケーションから—, 情報処理学会研究報告, 2013-GN-89(13), pp.1-6 (2013).
- 7) 服部哲, 柴田邦臣: 聴覚障害児・者のコミュニケーションを支援するAndroidアプリの開発, GNワークショップ2014論文集,

pp.1-6 (2014).

- 8) 服部哲, 柴田邦臣: 音声認識を用いたスケジューラの開発, インタラクシオン 2015, A64 (2015).
- 9) 中園薫, 織田修平: 聴覚障害者支援技術研究のレビューと将来への展望, 電子情報通信学会技術研究報告, Vol.109, No.358, pp.65-72 (2010).
- 10) 河原達也: ICT・音声認識の活用による講演・講義の字幕付与, 情報処理, Vol.56, No.6, pp.543-546 (2015).
- 11) 廣富哲也, 田中教子, 稲村駿: 広汎性発達障害児・者の会話理解を支援するコミュニケーションエイド, 電子情報通信学会論文誌. D, 情報・システム, Vol.97, No.1, pp.117-125 (2014).
- 12) 中村雅巳, 安田清: 聴覚障害者のための音声認識による日常会話支援機器の開発, ヒューマンインタフェース学会研究報告集, Vol.14, No.2, pp.37-38 (2012).
- 13) 鳥羽祐輔, 堀内大祥, 松本真佑, 佐伯幸郎, 中村匡秀, 内野智仁, 横山知弘, 武林靖浩: 聴覚障害者支援のためのマルチモーダル発話可視化に関する研究, 電子情報通信学会技術報告, Vol.114, No.500, pp.191-196 (2015).
- 14) Torii, I., Niwa, T., Ohtani, K., Ishii, N.: Study and Development of Schedule Application for Autistic Children, IIAI International Conference on Advanced Applied Informatics (IIAIAAI), pp.251-255 (2013).
- 15) 銭神裕宜, 小西遼, 清田公保, 合志和洋, 藤澤和子: 動的 Web ベース・データアクセスによる視覚シンボルを用いたユビキタスコミュニケーション支援システムの開発, 情報処理学会研究報告, Vol.2009-HCI-135, No.6, pp.1-6 (2009).