

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

服部 哲^{†1} 柴田邦臣^{†2}

筆者らは、聴覚障害児・者のコミュニケーション全般の支援を目標にタブレット・アプリの研究開発を進めている。本稿では、(1)リアルな会話そのものをサポートする機能、(2)コミュニケーションの結果を記録する機能、(3)コミュニケーションの記録を活用する機能を有する Android アプリについて詳細に述べる。本アプリは Android の音声認識機能を利用して、アプリ利用者（聴者）が発した言葉を音声認識し、その結果として取得できた文字列に対応する画像（ピクト）を検索し、アプリの画面に次々と表示する。会話の内容に合わせたピクトを表示することによって、聴覚障害者・児がその場の状況を理解することを可能にする。これまでのところ、本アプリの基本的な機能は実装済みであり、各機能の動作確認をおこなった。しかし、コミュニケーション全般を支援するには、音声認識エンジンの性能、認識結果の編集、記録した会話の活用方法などさまざまな課題が存在するため、実践的な評価を積み重ね、アプリの機能強化・拡張が必要である。

Android Application for Supporting Communication of Children and Persons with Hearing-Hard and Deaf

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

We have developed a communication assist software for children and persons with Hearing-Hard and Deaf using tablet media. In this paper, we discuss the Android application. It has the following three functions: (1) to support real conversations, (2) to record the communication results, (3) to make use of the records. Our Android application captions the real conversation through the speech recognition function of Android, and showing the photographs and pictograms corresponding to the captions to support the communication for children and persons with Hearing-Hard and Deaf. So far, we have implemented the basic functions of our application and tested them. As the results, we recognized its potential and problems.

1. はじめに

新生児聴覚スクリーニングが開始され、早期に難聴が発見されるケースが増加している[1]。日本産婦人科医会の調査報告によると、いくつかの問題点が指摘されているものの、2013年度の調査時点では分娩取扱機関の88%において新生児聴覚スクリーニング検査が実施されている[2]。新生児期に発見される早期療育が必要な聴覚障害の頻度は1000出生中1~2人といわれている。しかし、早期療育がおこなわれれば、コミュニケーションの発生、言語の発達・獲得につながり、生活の質の向上が期待される。一方、厚生労働省の平成18年身体障害児・者実態調査によると、日本の聴覚障害者数は18歳以上で276000人、18歳未満で15800人であり、1000人に2~3人が聴覚障害者であるという結果となっている[3]。

聴覚障害は音声情報の取得やコミュニケーションの障害とも言われており、その音声コミュニケーションの困難を補う、または代替するために、手話や筆談、絵カード、写真など、さまざまな補助手段・工夫が用いられている。

絵カードや写真は聴覚障害児・者の言語指導の一手法としても用いられている。このようなコミュニケーションの補助手段・方法は AAC (Augmentative and alternative communication) と呼ばれており[4]、筆者らも AAC に着目している。特に非音声系の補助手段・工夫である絵カードや写真に着目し、AAC をさらにコミュニケーション全般に適用できないかと考え、研究開発を進めてきた。

具体的には、聴覚障害児・者のコミュニケーション支援の助けとなるために、

- (1)リアルな会話そのものをサポートする機能
- (2)コミュニケーションの結果を記録する機能
- (3)コミュニケーションの記録を活用する機能

これらの機能を有する、Android タブレット上で動作するアプリを開発している[5][6]。

本アプリは、会話の内容を常に音声認識し、その場面に合わせて絵や写真（本研究ではピクトと呼ぶ）を表示することによって、会話の内容の理解とそれにもとづくコミュニケーションを支援する。また、会話を音声認識した結果として取得された、一連の文字列とピクトを保存し、ライフログのように活用できるようにする。

コミュニケーションが成立するには、互いの状況や文脈

^{†1} 駒澤大学
Komazawa University

^{†2} 津田塾大学
Tsuda College

を共有する必要があり、本アプリでは(1)の機能により、状況理解を支援する。また、今日ではブログや Twitter, LINE などさまざまなコミュニケーションツールが開発され、コミュニケーションは対面だけでなく、ネット上でもおこなわれる。本アプリでは(2)や(3)の機能により、対面以外のコミュニケーションをも支援する。

これまでもピクトを利用して聴覚障害児・者のコミュニケーションを支援することはおこなわれている[7]。また最近では、コミュニケーション支援の土台としてスマートフォンやタブレットが注目されており、それらの端末で動作するアプリも研究開発されている。たとえば、「こえとら」では聴覚障害児・者がキーボードから文字を入力すると音声で読み上げたり、聴者が音声入力した結果を画面に文字で表示したりすることができる[8]。また、「vocaco」や「トーキングエイド for iPhone/iPad」ではピクトを利用してメッセージを組み立てて音声で読み上げることができる[9][10]。これらのアプリは主に、聴覚障害児・者が知りたいこと伝え、理解することを目的としている。本研究のように、リアルな会話そのものをサポートし、コミュニケーションの記録の蓄積と活用をおこなうような、聴覚障害児・者のコミュニケーション全般を支援しようとする試みは筆者らの知る限りおこなわれていない。本稿では筆者らが開発している Android アプリの概要と詳細を述べる。

2. 本アプリの概要

2.1 アプリの特徴

本アプリは Android タブレット上で動作する。本アプリは Android の音声認識機能を利用して、本アプリの利用者（聴者）が発した言葉を音声認識し、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、端末の画面に次々と表示する。図1のように、会話の内容に合わせたピクトを表示することによって、聴覚障害児・者がその場の状況を理解することを可能にし、彼ら／彼女らのコミュニケーションを支援する。これにより、(1)リアルな会話そのものをサポートする機能を実現する。

また本アプリでは、会話を音声認識した結果として取得された、一連の文字列とピクトをシーケンスとして保存することができる。そして、保存されたシーケンスを適宜読み込み、そこに含まれる文字列とピクトの並びを再現することができる。また、ピクトの並びを自由に変更し保存しなおすこともできる。シーケンスをブログや Twitter などのソーシャル・メディアに投稿する仕組みは今後の課題であるが、これらによって、(2)コミュニケーションの結果を記録する機能と(3)コミュニケーションの記録を活用する機能を実現する。

2.2 アプリの構成

図2に本アプリの基本構成を示す。本アプリは3つの機能モジュールから構成される。それらはピクト登録、シーケンス作成、シーケンス閲覧・編集である[6]。カメラと音声認識については Android タブレットにあらかじめインストールされているアプリや機能を利用する。データベースはピクト管理用とシーケンス管理用のそれぞれ2つずつ、計4つのテーブルから構成される。ピクト管理用のテーブルは、ピクトの画像ファイルを管理するテーブルとピクトに付与されたタグ（キーワード）を管理するテーブルである。ピクトの検索では音声認識の結果とタグとを比較する。シーケンス管理用のテーブルは、シーケンスのタイトルや作成日時などを管理するためのテーブルとシーケンスに含まれるピクトとその並び順を管理するためのテーブルである。

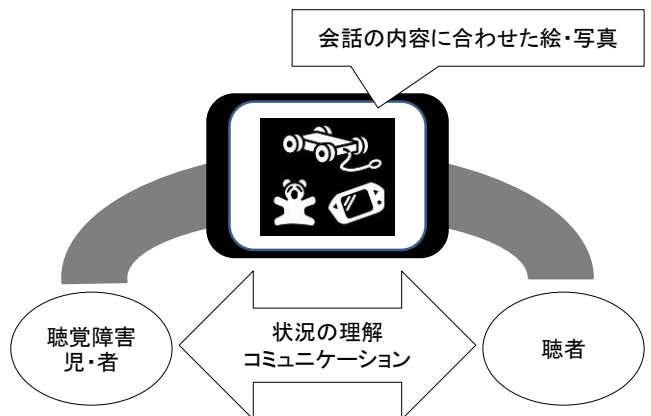


図1 本アプリの概念図

Figure 1 Concept of our application

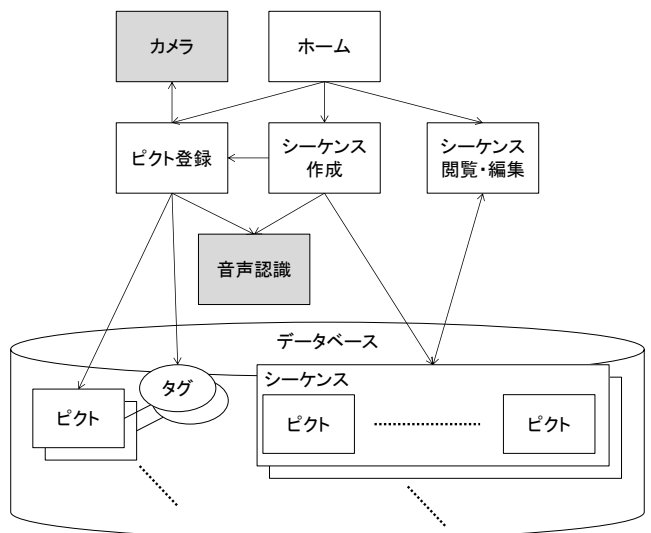


図2 本アプリの構成

Figure 2 System structure of our application

2.3 ユーザインタフェース

図3に本アプリの画面構成を示す。本アプリの画面は大きく4つの領域に分かれている。それらは①コンソール領域、②ピクト表示領域、③テキスト表示領域、④機能領域である。画面右上(図3中の①)はコンソール領域である。コンソール領域は1つのテキストビューから成り、そのビューをタップすることによって本アプリの機能を実行することができる。画面の一番大きな領域(同②)はピクト表示領域である。そこには音声認識に従ってピクトが並べられたり、作成したシーケンスが表示されたりする。シーケンスは左右にスクロール可能である。カメラアプリを起動したときはピクト表示領域にカメラによって取得した画像が表示され、その画像をタップすることによって写真画像が保存される。ピクト表示領域の下(同③)はテキスト表示領域であり、テキストビューが並んでいる。それらのテキストビュー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。直近の認識結果がテキスト表示領域の中央になるように自動的にスクロールされる。画面の右側や左側には、ピクトの検索や登録を実行するための矩形領域が存在する(同④)。音声認識の結果を表示するテキストビューをタップするとそのコピーが作成され、それを矩形領域にスワイプすることによって、ピクトを検索したり、ピクトにタグ付けしたりすることができる。それらの機能を実行するときのみ機能領域が表示されるようになっている。コピーされたテキストビューはピクトの検索やタグ付けがおこなわれると消去される。

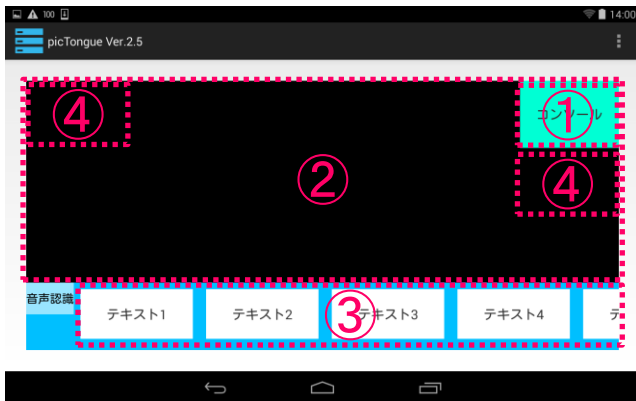


図3 本アプリの画面構成

Figure 3 User interface.

3. アプリの詳細

3.1 音声認識

Androidアプリで音声認識機能を利用する方法としては、Androidタブレットの音声認識機能を利用する方法や既存の音声認識エンジンを組み込む方法[11]などがある。本研究では、まずはアプリの試作を優先しているため、Android

タブレットの音声認識機能を利用することにした。

Androidタブレットの音声認識機能をアプリから利用する場合、

- Android標準の音声認識ダイアログを呼び出す
- `SpeechRecognizer` オブジェクトを利用して、バックグラウンドで音声認識機能を実行する

という2つの方法がある[12].

本研究は言語獲得とコミュニケーション支援のために「状況を定義」することから始めた。そのためには、話者(聴者)が「何について話すか」を示すことができるようにする必要があり、初期の実装では、Android標準の音声認識ダイアログを呼び出して音声認識をおこなった。

しかし、リアルな会話そのものをサポートするには、音声認識をおこなうたびにダイアログを呼び出すのではなく、コミュニケーションと同時進行的に会話を音声認識していく必要がある。そのため、本研究では `SpeechRecognizer` オブジェクトを利用する方法も実装している。

本アプリは音声認識の結果として取得された文字列を一つひとつテキストビューに設定することによって記録する。

3.2 ピクトの検索

本アプリのピクト検索はシーケンス作成モジュールが起動されているときに実行されるが、その実行は2つのタイミングでおこなわれる。

ひとつは、音声認識の結果として文字列が取得されたときである。この場合、その文字列がタグ付けされているピクトが検索される。もうひとつは、テキスト表示領域のテキストビューを、ピクト検索の機能領域へスワイプしたときである。この場合、テキストビューに表示されている文字列がタグとし付与されているピクトが検索される。これらの方法により会話の内容の理解とそれにもとづくコミュニケーションを支援する。

後述のピクト登録機能によって、本アプリでは同じタグを複数のピクトに付与することができるため、音声認識の結果、あるいはテキストビューをスワイプして特定の方向へ移動させた結果として検索されたピクトを別のものに変更したいこともある。本アプリでは、検索されたピクトを長押しすることによって、そのピクトと同じタグが付与されたピクトをすべて表示することができる。そして、その中から利用者がピクトを選択することによって、そのピクトがそれまで表示されていたピクトの代わりに画面上に表示されるようになる。本アプリでは、ピクトの検索結果の優先順位を管理しており、利用者がこのピクト変更操作を一度おこなうと、本アプリは日本語入力システムにおける変換候補の優先順位のように、選択されたピクトの優先順

位が一番高くなるように順位を変更する。したがって、それ以降は同じタグによってピクトが検索されたとき、利用者が選択したピクト、つまり優先順位が一番高いピクトが検索されるようになる。

3.3 シーケンスの作成・閲覧・編集

本アプリでは、音声認識の結果として取得された一連の文字列と画面に表示されたピクトは、シーケンスとして保存される。シーケンスの保存時にはシーケンスが保存された日時と任意で位置情報も記録する。そして本アプリでは、保存されたシーケンスを読み込み、文字列とピクトの並びを再現することができる。そのため、会話の記録・コミュニケーションのログとなり、利用者は後でその内容を確認したりブログや Twitter などのソーシャル・メディアで活用したりすることができる。スマートフォンなどに搭載されている各種センサーによって自動的に記録されるライフログと違い、本アプリでは利用者自身が残したい、あるいは、残すべきと考えたものだけを記録することができる。また本アプリでは、ピクトの並び順を自由に変更することができるようになってきている。この機能によって、利用者は、時系列に並んだピクトを、その会話の内容を理解しやすいように並べ替えて活用することができる。ただし、ソーシャル・メディアとの連携など記録されたシーケンスを具体的にどのように活用するのかについては今後の課題である。

3.4 テキストビューのスイプ

本アプリでは、ピクトを検索するときや、新しいピクトにタグ付けをおこなって登録する場合、利用者はテキスト表示領域のテキストビューをそれらの機能を実行するための機能領域の方向へスイプすればよい。

具体的には、本アプリの画面左上や右上には各機能と結び付けられた、非表示設定のテキストビュー（スイプするテキストビューと区別するために機能領域ビューと呼ぶことにする）が配置されている。利用者がそれらの機能領域ビューに向かってテキストビューをスイプさせることによって、そのテキストビューが機能領域ビューに向かって移動するようにアニメーションがおこなわれ、アニメーション終了後に機能領域ビューに結び付けられた機能が実行されるようになってきている。どの機能領域ビューに向かっているかは、スイプ操作の方向を随時追跡し、その方向を示すベクトルと、スイプ中のタッチイベントの発生位置から機能領域ビューの中心点へと向かうベクトルとの類似度によって求めている。タグや検索用キーワードはスイプされたテキストビューから取得される。

筆者らは、本アプリの機能をそれぞれ別のクラスとして実装しているため、利用者からのフィードバックを踏まえての機能の拡充や変更をおこないやすくしている。現在では本アプリの機能をカスタマイズすることはできないが、

ピクトの検索やタグ付け以外の機能を利用者が自由に選択可能にすることによって、本アプリを利用者のニーズに合わせてカスタマイズできると考えている。

3.5 ピクト登録

聴覚障害児・者の言語教育の分野では、言語指導の一手法として絵カードや写真が用いられており、本研究はその手法にヒントを得ている。そのため本アプリでも、「コミュニケーション支援用絵記号デザイン原則 (JIS T0103)」に記載されている約 300 の絵記号例[13]を利用してピクトを用意し、インストールと同時に利用可能な状態となる。しかし、会話の場面に合わせたピクトをすべてあらかじめ用意しておくことは現実的ではない。そのため、本アプリでは、Android タブレットのカメラアプリを利用してその場で撮影した写真や、あらかじめ撮影しておいた写真を選択し、その写真にタグを付与することによって、利用者が自由にピクトを追加できるようになっている。また、1つのピクトに1つのタグしか付与できない場合、コミュニケーションの豊かさが疎外されてしまう。そのため、データベースのテーブルを工夫することによって、同じタグを複数の写真に付与したり、同じ写真に複数のタグを付与したりすることも可能になっている。したがって、コミュニケーションの場面に適したピクトをその場で追加して利用することができる。

なお、これまでに述べた機能のほかに、データベースの各テーブルを CSV 形式のファイルとして SD カードにバックアップすることができる。

3.6 実装

筆者らはこれまでに上記で述べた機能を実装し、複数のタブレットにインストールして動作確認をおこなった。開発環境は Eclipse であり、Android OS のバージョン 4.2 や 4.4 のタブレットにインストールした。アプリを構成する Java のクラスファイル数は 13、レイアウト用 XML ファイルの数は 10 である。テキストビューをスイプすることによって Twitter に投稿する機能も実装したが、アプリには組み込んでいない。

図 4 はシーケンス作成のスクリーンショットである。ピクト表示領域には、図 4 中の①アプリのインストール時に取り込まれるピクト（黒の背景に白で描かれた絵）と、同②利用者が登録したピクト（本の写真）が表示されている。テキスト表示領域の緑色のテキストビュー（同③）はコピーされたビューであり、このビューをスイプすることが可能である。

図 5 はシーケンス編集のスクリーンショットである。シーケンスの閲覧時はシーケンス作成時と同じように、ピクト表示領域にピクトが並ぶようになっているが、シーケン

ス編集時はピクトと対応する文字列のセットを1つのペアとしてそれらがリスト表示される(図5中の①)。シーケンス編集時は、順番を変更したいペアを目的的位置まで上下にドラッグすることによってシーケンス内のピクトの順番を変更することができる。図5では、②の「勉強する」というピクトの順番を変更しているところである。



図4 シーケンス作成
Figure 4 Making sequence.



図5 シーケンス編集
Figure 5 Editing sequence.

図6は新しくピクトを登録する画面のスクリーンショットである。本アプリでは、その場で撮影した写真やあらかじめ撮影しておいた写真をピクトとして利用可能であるが、図6はその場で撮影した写真にタグ付けし、ピクト登録をおこなっているところを示している。その場で写真を撮影する場合、カメラプレビューがピクト表示領域に表示されるため、利用者はそのプレビューをタップすることによって、写真を撮影することができる。撮影された写真は図6中の①のようにピクト表示領域に表示されるため、その写真で問題なければ、利用者は、タグが表示されているテキストビューをコピーして、そのコピーをタグ付け登録の機能領域の方向へスワイプする。図6中の②のテキストビューはコピーされたビューである。同③のように、本アプリ

では、テキストビューを機能領域の方向へスワイプすると、その領域が表示されるようになっている。



図6 ピクト登録
Figure 6 Adding pict.

4. 考察

絵や写真によるコミュニケーションをおこなうには練習が必要である[14]ものの、「何について話すのか」を示すことによって、コミュニケーション場面の状況を把握し、会話の内容への関心が高まり、そして理解しやすくなるのであれば、本アプリは聴覚障害児・者のコミュニケーション支援として十分に可能性があるのではないだろうか。また、写真にタグをつけて整理するという本アプリのアイデアはタブレットにおける従来のファイル管理方法と異なり、多様な視点で写真を整理することになり、その機能だけでも価値があると思われる。

しかし、コミュニケーションは双方向のやり取りがあって成立するものである。聴者の話を音声認識しピクトを次々と表示することはコミュニケーション場面の状況把握の手助けになる一方、コミュニケーションの双方向性の一方を支援しているのみである。現状では、本アプリは聴覚障害児・者が相手の意見を受け取り自分の考えや思いを表出することを支援していない。ピクトを並べることによって自らの考えや思いを伝えられるような仕組みが必要である。その場合、いかにピクトを利用した表出方法を習得するかも並行して検討しなければならない。

本アプリの核となる技術の1つは音声認識である。50年以上にわたる音声認識技術の研究開発によって、今日では、カーナビやスマートフォンの音声アプリなど音声認識技術が身近になりつつあり、その性能も高くなっている。しかしながら現状の技術では、日常会話や雑音の多い環境下の認識では、その性能は大きく落ちる[15]。本アプリにおいても、Android標準の音声認識ダイアログを表示して、それに向かって「ていねい」に、かつ「はっきり」と発声すれば、かなり高い精度で認識し、的確にピクトを検索する

ことができる。しかし、SpeechRecognizer オブジェクトを利用してバックグラウンドで音声認識機能を実行する場合、そのオブジェクトの特性を把握して発声しなければならないが、それよりも問題なのは、ほとんど認識できないことである。本アプリの音声認識はクリアすべき大きな課題である。音声認識エンジンを利用者に合わせてカスタマイズしていく必要がある。

音声認識の結果が誤っていたり、利用者が言い間違いをしたりするという事は生じるものである。そのためには、シーケンス中の文字列を修正可能にしなければならない。現状では、本アプリのシーケンス編集機能はピクトの並び順を変更することのみ可能であるため、たとえば、テキストビューを長押しすることによって、そのテキストを修正可能な状態にするなど、シーケンス編集機能を強化しなければならない。また、シーケンスを編集した結果を、テキスト表示領域のテキストビューをスワイプするような簡単な方法でソーシャル・メディアに投稿可能にするなど、シーケンス編集から活用までをつなぐような仕組みが必要であろう。

一方で、記録したシーケンスをどのように活用していくのかも十分に検討していかなければならない。ソーシャル・メディアの WebAPI を利用すれば、シーケンスをソーシャル・メディアに投稿することは可能であるが、それをおこなうことの意義や効果をしっかりと検討し、本アプリで記録したシーケンスをどのように活用していくのかは今後の課題である。筆者らは聴覚障害児・者と聴者とのコミュニケーションに注目しているが、聴覚障害児・者と聴者とのコミュニケーション場面だけでなく、たとえば、聴者同士の会話を記録し、聴覚障害児・者と会話の内容を共有するという活用は1つのあり方として考えられるのではないだろうか。聴者である子どもが経験したことを本アプリで記録し保護者と共有するという活用は1つのあり方として考えられる。

5. おわりに

本稿では、聴覚障害児・者のコミュニケーション全般の支援を目標に、(1)リアルな会話そのものをサポートする機能、(2)コミュニケーションの結果を記録する機能、(3)コミュニケーションの記録を活用する機能を有するアプリについて詳細に述べた。現在、アプリの基本的な機能は実装済みである。実装機能の動作確認をおこなう中で機能面の課題が明らかになった。たとえば、音声認識の精度、シーケンスの編集である。また、記録したシーケンスをどのように活用するのかという応用面かも検討しなければならない。アプリの操作性や性能などの評価も残されている。

今後は、これらの課題を解決するために本アプリの開発を進めるとともに、福祉現場のニーズの把握や理論的な研

究をおこない、本アプリを福祉の現場に導入し実践的に評価しなければならない。それによって、真に聴覚障害児・者に寄り添うようなコミュニケーション支援アプリの開発を目指したい。

参考文献

- 1) 大迫茂人: 第4章 新生児聴覚スクリーニングの現状と課題, ろう教育学会編「聴覚障害教育の歴史と展望」, 風間書房, pp.107-130 (2012).
- 2) 日本産婦人科医会: 新生児聴覚検査の実態調査報告 (2014). http://www.jaog.or.jp/all/document/77_140611.pdf
- 3) 厚生労働省: 平成18年身体障害児・者実態調査結果 (2008). <http://www.mhlw.go.jp/toukei/saikin/hw/shintai/06/dl/01.pdf>
- 4) 中邑賢龍: AAC 入門, ころりソースブック出版会 (2014).
- 5) 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, Vol.2013-GN-89, No.13, pp.1-6 (2013).
- 6) 柴田邦臣, 阿由葉大生, 小河千了, 服部哲: タブレット・メディアによる「拡張コミュニケーション」—聴覚障害児に学ぶライブログの活用—, GN ワークショップ 2013 論文集, pp.1-2 (2013).
- 7) 金森克浩: 特別支援教育における 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).
- 8) 情報通信研究機構: こえとら. <http://www2.nict.go.jp/univ-com/plan/applications/koetra/>
- 9) Yuuichi Kaga: vocaco. <https://itunes.apple.com/jp/app/vocaco/id744835195?mt=8>
- 10) ユーブラス: トーキングエイド for iPhone/iPad. <http://www.talkingaid.net/products>
- 11) 山本大介, 大浦圭一郎, 西村良太, 打矢隆弘, 内匠逸, 李晃伸, 徳田恵一: スマートフォン単体で動作する音声対話 3D エージェント「スマートメイちゃん」の開発, インタラクション 2013, IPSJ Symposium Series Vol. 2013, No.1, pp.675-680 (2013).
- 12) 高見知英, 菅野祥礼, 神原健一, 茶園亮, 松岡謙治: Android API プログラミング・リファレンス, マイナビ (2012).
- 13) 共用品推進機構: 「コミュニケーション支援用絵記号デザイン原則 (JIS T0103)」に収載されている絵記号例. http://www.kyoyohin.org/ja/research/japan/jis_t0103.php
- 14) 藤澤和子, 井上智義, 清水寛之, 高橋雅延: 視覚シンボルによるコミュニケーション: 日本版 PIC, プレーン出版 (1995).
- 15) 河原達也: 音声認識技術の現状と将来展望, 電気学会誌, Vol.133, No.6, pp.364-367 (2013).