

OTON GLASS を利用したディスレクシアの人のための音声読み上げ機能拡張システムの評価

應武双葉^{†1} 栗原一貴^{†1}

概要: 本論文では, 学習障害の1つであるディスレクシアと呼ばれる障害を持つ人々に対し, 音声合成技術を用いて日常生活を支援していくためのシステムの提案を行う. メガネ型デバイス OTON GLASS は, カメラが捉えた画像を認識し, その人の視界にある文字が音声合成によって読み上げられるシステムである. 我々は OTON GLASS に音声読み上げに関するパラメータ調節の拡張機能を実装した. 評価実験により有効な点, また改善すべき点が得られた.

An Evaluation of Extended OTONGLOSS with Regard to Speech Synthesis Parameters for Supporting Dyslexics

FUTABA OTAKE^{†1} KAZUTAKA KURIHARA^{†1}

Abstract: In this paper, we propose a system to support daily living using speech synthesis technology for people with disabilities called dyslexia which is one of learning disabilities. OTON GLASS is a glasses device system in which an image captured by a camera is recognized, and characters in the view of the person are read out by voice synthesis. Using the glasses device, we have implemented the extended functionalities with regard to speech synthesis parameters. An evaluation shows the effectiveness and implications for future improvement.

1. はじめに

2016年4月に障害を理由とする差別の解消の推進に関する法律(以下障害者差別解消法[1])が施行されて以来, 障害者に対する社会的な援助や後押しは急務となりつつある. 情報分野も同様であり, 特に技術的な支援においては大きな可能性を持つものも多い. 学習障害の1つであるディスレクシアもその1つである. 日本人の3~5%程度が該当する[2]と言われているこのディスレクシア障害を持つ人たちは, 知的能力や一般理解能力などに特に異常がないにもかかわらず, 文字言語の認知に著しい困難を抱えている.

ディスレクシア障害に関しては一般的に情報技術による補助が有効な場合が多々ある. これは音声読み上げなどによって「読む」という作業を補助することが可能であるからだ. 実際に学習を補助するための読み上げ補助システムとして「DAISY [3]」などがあげられる. WEB上のドキュメント読み上げには「ドキュメントトーカー [4]」などもある. ところが本障害は非常に認知度が低く, 具体的な支援に手が届いていない部分が多い.

そこで我々は事前調査と題し, ディスレクシアを持つ女子高校生1名, 保護者7名に話を聞いた. その結果大きく2つの意見が得られた. 1つは日常において音声読み上げがないという点で多々困難を抱えているということである. 中には音声読み上げを行ったことでテストの点が20点上がったという報告もあった. もう1つは, 個人によって聞

き取りやすい音声に差異があるということである. 例えば男性の声は気持ちが悪くて聞くことができない, 声の高い音は聞き取れないなどがこれにあたる. 現状の音声読み上げ技術では選べる読み上げの音は一律か, 存在しても男女の性差だけという場合が多い. 従ってただ音声読み上げを利用するだけではなく, 声の高さ, 速さを変えてより個人の聞き取りやすく快適な音にカスタマイズすることが必要なのではないか, という知見が得られた.

本研究ではこの事前調査を踏まえメガネ型デバイス OTON GLASS に対し, 自由に声の高さ, 速さを操作できる拡張したシステムを実装した. OTON GLASS はメガネに搭載されたカメラで文字を読み取り, それを音声読み上げすることで視覚障害を持つ人をサポートするメガネである.

本稿ではこのシステムについての研究結果を述べる. それに先立ち, まず第2章で関連研究を概説する. 次に第3章でシステムの説明, 第4章ではこのシステムを用いて行った評価を述べ, 第5章で今後のまとめと考察, ディスレクシア支援に対する展望を述べるものとする.

2. 関連研究

2.1 ディスレクシア障害への支援

ディスレクシア支援は, 近年徐々に取り組まれるようになってきた. ディスレクシアはその特性から特に教育過程で極めて著しい困難を抱えることが多い. 最も良くある支援の中では別室受験を認める, 問題文を教員が読み上げる,

^{†1} 津田塾大学
Tsuda University

板書を補助する、授業の録音、動画撮影を許可する、などといったことが報告されている[5].センター試験でも 2011 年度から発達障害やディスレクシアを含む LD を持つ子供達への配慮が行われきた[6].これらの支援はあくまで人力による手段を頼ったものであり、対人があることが前提にある。しかし自明なことであるが、常に誰かが補助できるとは限らないのが実情であり、そのためには文字や音そのものへの理解を深めることが重要である。本研究はディスレクシア障害を抱えた方の中でも読むことに対する支援に特化したものであるため、関連研究ではその研究を列挙する。

(1) Evaluation of DysWebxia[7]

文字そのものに対して行われた研究である。iOS デバイスを使ってディスレクシアの人を支援するアプリを用い、評価を行ったものである。テキストを読み込み、それを翻訳して類義語などを表示する。この際にフォントサイズ、色、フォントなどを変えるなどしてフォント そのものを見やすくすることができる。評価の結果この手法は有効であったと示唆された。視覚情報を補うツールとしての研究は他に WEB デザインを変える Firefixia[8], 実際にどのフォントサイズが見やすいかを調べた研究[9], 見やすいフォントを調べた研究[10]などがあげられる。本研究では文字そのものではなく、読み上げて文字情報を補おうとしている点が異なる。

(2) Assistive systems[11]

音声を明瞭化することによってディスレクシアの人への支援の行い方を考えた研究である。母音と子音の比率、母音の長さ、間の取り方、話す速度や声の高さなどを変えることにより聞き取りやすさの改善を提案している。この研究に対して、本研究では読み上げ速度、声の高さなどの変更による聞き取りやすさの検証を行っており、整合性のある結果が得られている。

(3) DAISY 風テキストリーダー[12]

ディスレクシア障害で対象となる児童生徒への学習的支援は、現状では DAISY 教材を用いることが多い。しかしながらこれを日常的に作成するには複雑な手順を踏む必要があるため決して容易とはいえない。DAISY 風テキストリーダーはこの機能をできるだけ容易に使用するために、テキストファイルを作成するだけのおおむねほぼ DAISY と同じような表示および読み上げを行うことができたようにしたのである。

2.2 メガネ型デバイスによる支援

英語圏では OTON GLASS と似たようなシステムは既に考案されている。OrCam[13]は 2010 年にイスラエルのオラカム社が弱視を持つ人向けに開発したガジェットである。小型カメラを眼鏡につけることにより、目の前で指差したものを読み上げることができる。言語は読み上げ、認識ともに英語に限る。読み上げ音声そのものを変更させるとい

う機能はない点が本研究との差異である。

3. システム

本研究では声のパラメータを変更することはディスレクシアの人によって有効である。ということへのさらなる裏付けと、パラメータ変更時のレスポンスについて考えることを主としている。したがって、パラメータ変更はより使用者にわかりやすいように、本体に組み込むのではなく、比較的認知度が高く使用しやすいファミコンのコントローラで行うものとした。

3.1 基本設計

ウェアラブルデバイスそのものは OTON GLASS から踏襲するものとした。本体での言語は Python を使用している。そこでシステム本体に前述のファミコンのコントローラを利用し、これによりパラメータを調節できる機能を設定した。使用した機材は図 1 の通りである。



図 1 使用した機器

Figure 1 The system overview.

3.2 パラメータ調節

声の高さ、速さを変更できるボタンについてコントローラを使って実装した。また読み上げ中の一時停止、再生、1秒ずつの早送り、巻き戻し、現在の声のパラメータについて確認ができるボタンも付け加えた。実際のコントローラの機能を説明したものが図 2 である。

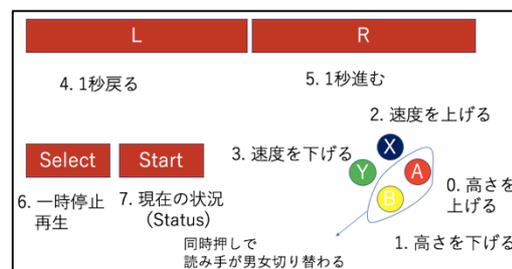


図 2 コントローラの機能

Figure 2 The functionalities of the controller

声の高さは-50 から 50 まで 10 刻みで 11 段階を調節する

ことができる。同様に声のスピードについて、「とても速い」「速い」「普通」「ゆっくり」「とてもゆっくり」の5段階がある。これは音声読み上げを行う AWS の Amazon Polly[14]の仕様によるものである。

4. 評価

2017年12月4日に、拡張した OTON GLASS で音声読み上げを行うことで、実際に理解度が高まるのか、という評価を行った。また同様に音声の高さ、速さを変える需要がどれほど存在するか、変更するのであれば操作のフィードバックをどのように行うべきか、またこのシステムが当事者にとって心地よいものであったか、という評価も実施した。

4.1 評価協力者

ディスレクシアを持つ年長から 中学 1 年生までの児童生徒 12 名 (男 11 : 女 1) を対象に行った。なお、調査に当たっては協力者が未成年であることを受け、学習障害 (LD) を含む子供の支援団体 (ソラアル SSE) に協力を要請し、当該児童生徒の保護者に事前に説明の上、許諾をとり実施した。表 1 に協力者の詳細情報を示す。

表 1 評価協力者

Table 1 The participants' profiles.

ID	1	2	3	4	5	6
パターン	I	II	III	IV	I	II
学年	中1	年長	年長	小3	小6	小3
性別	男	男	男	女	男	男
ID	7	8	9	10	11	12
パターン	III	IV	I	II	III	IV
学年	小1	小3	小6	小1	中1	中1
性別	男	男	男	男	男	男

4.2 評価方法

小学 1 年生学習程度のひらがなと漢字で構成された 3 文程度の 2 つの文章 A, B を用意した。今回の評価には年長に当たる幼児も混じっているが、この幼児らに関しては事前に 1 年生程度の漢字であれば読めることを確認した。

各文章を読むに際して、まずは読み上げがあるかないかの順番を入れ替えた表 2 のように 4 つのグループに分類する。さらにこの 4 グループを表 3 のように 4 パターンに組み合わせる。これに伴って、一人当たり 2 つの文章を 2 パターンずつ総計 4 回読み取りすることになった。なお順序を入れ替えたのは、1 度文章を読んだ後の順序効果と学習効果を考慮し、これを相殺するためである。評価時間に関しては協力者の中に 10 分以上の継続した作業に苦痛を感じる子供がいたことを配慮し、文章も健常児であれば長くても 30 秒程度で読めるものを選んだ。また各文

章を読む度に、「文章の内容が理解できたかどうか」の程度について、パソコン上のヴィジュアルアナログスケール (VAS) で評価した。また 4 グループそれぞれが終わった時点で「快適に拡張した OTON GLASS が利用できたか」について上記と同様に VAS を用いて評価した。VAS の値は最高点 1、最低点 0 に規格化し集計した。問いに関しては口頭で読み上げを行い、評価項目で印をつける作業以外に協力者から何かコメントが得られた場合にはこれを代筆した。

さらにこれとは別に、スピード調節と声の高さにおいて個人差が生じるのではないかという事前調査を踏まえ、評価を始める前に実際に評価協力者に聞き取りやすい値に調節してもらった。なおこの際にこの 2 点のパラメータを変更する時のレスポンスについても評価した。評価については、ボタンを押すと 3 パターンのレスポンスが変わる状況を作り上げ、これを 3 択にして選んでもらうこととした。1 つは「音声読み上げがボタンを押すたびに現在の状況を読み上げる」、1 つは「ボタンを押すたびに効果音が鳴り、最終的に調整した状況を読み上げる」、1 つは「ボタンを押すたびに LED で光が点滅し、最終的に調整した状況を読み上げる」ものである。音声は「速さ (高さ) は〇〇です。」という文章を読み上げた。〇〇の部分には操作時点での現状が入る。なお、効果音と LED についてはそれぞれ 2 パターンの音と光を用意した。これはパラメータを変える際に、上げる、下げるという対照的な操作があることから、差異を明らかに示すためである。

なお実際の評価の様子は図 3 の通りである。



図 3 評価の様子

Figure 3 A snapshot of the evaluation

表 2 組み合わせ表 1

Table 2 Task arrangement 1.

①	文章 A を読み上げあり▶なしの順で見る
②	文章 A を読み上げなし▶ありの順で見る
③	文章 B を読み上げあり▶なしの順で見る
④	文章 B を読み上げなし▶ありの順で見る

表3 組み合わせ表2

Table 3 Task arrangement 2.

I	①を読んだ後④を読む
II	②を読んだ後③を読む
III	④を読んだ後①を読む
IV	③を読んだ後②を読む

4.3 結果

(1) 理解度

まずは文章Aの読み上げあり、なし、文章Bの読み上げあり、なしの4つの文章を見た場合の評価の結果を評価協力者ごとにまとめたものが以下の図4である。0が最低、1が最高の理解度を表す。なお、質問では「非常に思う」に近い方が最高の理解度であり、「全くそう思わない」に近いほど最低の理解度である。図4の結果から、読み上げた条件の方が、ほとんどの評価協力者において高い数値を示した。

個別に具体例を記述する。IDが1,2,3,4,6,8,10だった評価者に関しては読み上げた方の理解度が高かった。IDが7,9の評価者は12人の中で特に著しくディスレクシアとしての困難があるものであり、とりわけ顕著な差が出ている。反対にIDが11,12の評価者は12人の中では軽い方の症状であり、それほど差がない結果になったのではないかと推測できる。IDが5の評価者はパターンIで評価したが、文章Bを読んだ後に「自分は人の声が好きで、音声読み上げだと棒読みのように聞こえ、聞き取りづらいことがある。」という旨の発言があった。

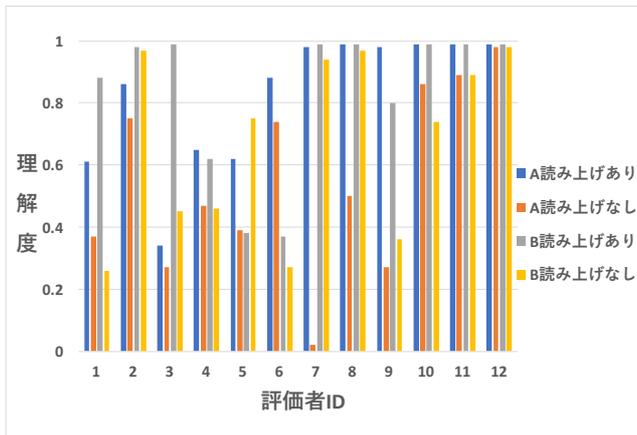


図4 協力者ごとの理解度

Figure 4 The levels of understanding of each participant.

次に読み上げあり、読み上げなしのそれぞれの分布の様子を図5で示した。対応ありt検定を用いて検定したところ、有意差が認められた ($p < 0.05$)。以上の結果から、ディスレクシアの人にとってメガネ型デバイスによる読み上げは有効であるという結果が示唆された。

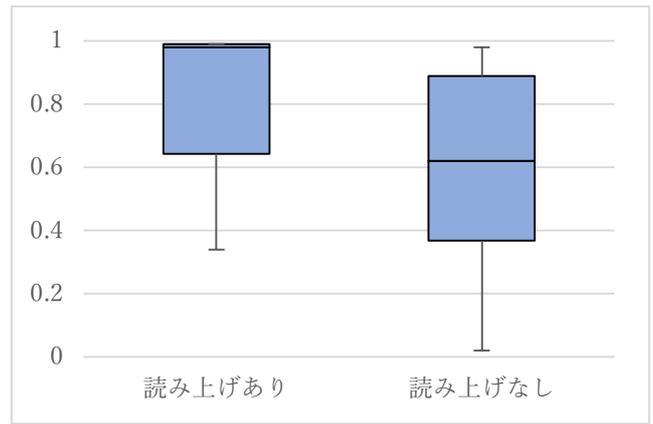


図5 読み上げの有無ごとの理解度の平均

Figure 5 The average levels of understanding based on the presence or absence of reading subtitles out.

(2) レスポンス

声のスピードや高さなどのパラメータを操作する際のレスポンスについて音声読み上げ、光、効果音の3通りを試してもらい最も良いと思うものを3択方式で答えてもらった。結果を表4で示す。この表により、音声読み上げは有効なものだと推察できる。具体的には、音声読み上げと答えたもののほとんどは「読み上げをした方が、内容確認がスムーズである」と答えたが、光と答えた3名について理由を問うと、単純に「目だけで分かる方がわかりやすい」と答えた協力者が1名、「音声で読み上げると文章が長く、その点光だとレスポンスが早い。」と答えた協力者が2名いた。うち1名は、短くなれば音声読み上げの方が良かったと答えていた。

表4 レスポンスの結果

Table 4 Preference with regard to the response modality

音声読み上げ	9
光	3
効果音	0

(3) 声の高さと速さ

音声読み上げのスピードと声の高さについて、協力者が評価に選んだこれらの値の設定を表5で示す。なおこれらのパラメータについて、デフォルトの値は2項目ともに1.0であり、コントローラにより調整するユーザーインターフェイスである (図2)。スピードについては、「とても速い(1.2)」「速い(1.1)」「普通(1.0)」「ゆっくり(0.9)」「とてもゆっくり(0.8)」として記載した。また、読み手の性別も変更可能としこれも記録した。デフォルト値は女であり、男をM、女をFとしている。

スピード調節をした協力者は12人中8人であり、67%がスピード調節を行った (図6)。また調節時の速さ

も 0.8 から 1.4 と多岐にわたり、ゆっくりにしたり速くしたりと人によって様々だということが分かった。

表 5 スピード、声の高さの調節の結果

Table 5 The proffered speed and pitch chosen by each participant.

ID	1	2	3	4	5	6
スピード	1.1	1.2	0.9	1	1	0.8
高さ	1.1	1	1.2	1	1.1	0.8
男女	M	M	F	F	M	M
ID	7	8	9	10	11	12
スピード	1.2	0.9	0.8	1.1	1	1
高さ	0.8	1	1.4	0.8	0.9	1
男女	M	F	F	M	M	F



図 6 スピード調節

Figure 6 Summary: the adjustments of playback speed.

次に声の高さについては図 7 の通りである。声の高さは 12 人中 8 人が変更しており、性別を変更した評価者も 7 名に及んだ。パラメータも「とてもゆっくり」から「とても速い」まで個人差があった。

従って声の高さ、速さを変えることでの、改めての有効性が示された。

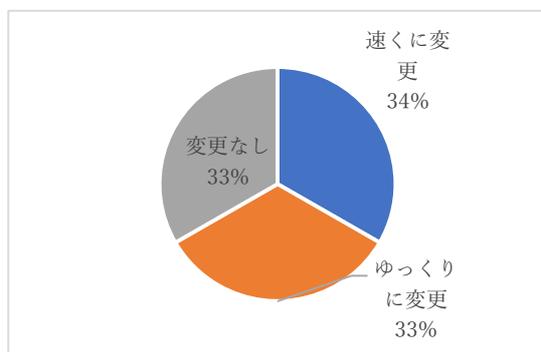


図 7 声の高さの調節

Figure 7 Summary: the adjustments of pitch.

(4) システムの快適性

本システムを使用することでの快適性を問うた結果を図 8 で示す。質問では「非常にそう思う」である 1 に近い

方が快適であり、「全くそう思わない」である 0 に近いほど不快を感じているものである。平均は 0.82、標準偏差値は 0.22 であり、メガネ型デバイス OTON GLASS に本人の好みの音声変更を加える機能を拡張したことによって高い数値を認めることができた。

ID5 の評価者については 0.5 を下回る結果になった。これは(1)理解度にも関連することであるが、本評価者はもともと人工音声よりも人間の声の方が聞き取りしやすいという特性があったことによるものである。

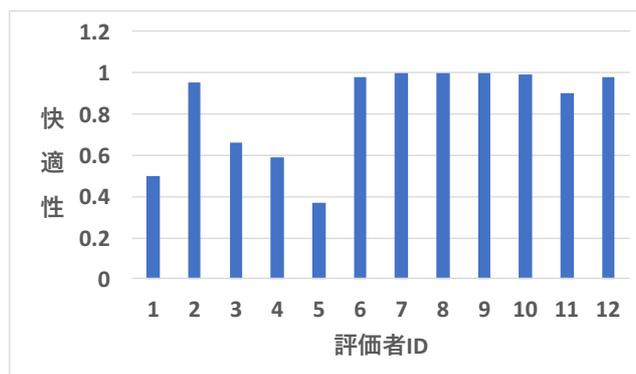


図 8 快適性

Figure8 Summary: degree of comfort.

5. 考察と今後の展望

本研究ではディスレクシア障害をもつ人たちが文字を見る際に補助をすることができるメガネ型デバイス OTON GLASS に音声の質を変更できる機能を拡張し、これについての評価を行った。その結果ディスレクシアの人にとって音声読み上げは非常に有効であり、個人に合わせて声の高さや速さを変えることができる必要があるということが改めてわかった。またこれは OTON GLASS 使用時だけではなく、声のパラメータ調節時の読み上げ機能についても同様である。

また今後のシステム上の課題として以下 2 点が挙げられた。

1 点目は、本追加点を実際にメガネ型デバイスに組み込むことである。本研究は声の高さ、速さという声のパラメータを変更することでより利用者のニーズに沿った有効的な手法を確立できたかを評価するためのものである。従って変更のためのインターフェイスを馴染みのあるコントローラで行っており、これはシステムとしては操作性に優れているとは言えない。ここで得られた知見をもとに、より良い方法での実装を行う必要がある。

2 点目に、個人の感覚の差を改めて意識する必要があるということである。ID5 の評価者など、音声読み上げがどうしても受け付けられないという人もある程度存在する。音声読み上げが多くの人にとって有効であることは示されてきたが、そうではない少数を切り捨てて

行くのは本意とは言えない。例えばボタン操作時のレスポンスなどに光と音声合成を組み合わせるなどの複合により、個人に寄り添ったより幅広いニーズに応えていく必要があると思われる。

本研究を通じて今後さらにディスレクシアの人に目が向けられ、支援研究が進展すれば幸いである。第1章でも述べたようにそもそもディスレクシアという障害は日本では認知度の低さもあり、存在そのものが程度の差があれ浸透していないという問題 [15]がある。認知度の低さは、時に当事者たちの生活を一層困難にさせる。特に音声読み上げへの理解がないことで日常生活に支障をきたしている、という意見は強く世間に訴えられるべきであると思われる。現状の社会ではまだ障害を持つ人たちへの援助という点では行き届いていないところも多い。特にディスレクシアのような障害に対しての理解が基盤になく、母数も少ない障害の場合はなおさらである。ディスレクシアの人たちになるべく不自由がない生活を送ることができるようになるためには社会の理解と、それを支える情報技術が不可欠である。今後の当該分野の研究の進展が望まれる。

謝辞 本研究はJSPS 科研費 JP15H02735,JP16H02867 の助成を受けたものです。また本研究を行うにあたり快く評価にご協力くださったソラアル SSE の皆様、株式会社 OTON GLASS の皆様に感謝いたします。

参考文献

- [1] “障害者差別法”.http://www8.cao.go.jp/shougai/suishin/law_h25-65.html,(参照 2017-12-05).
- [2] “通常の学級に在籍する特別な教育的支援を必要とする児童生徒の全国実態調査(内閣府 障害者)”.http://www8.cao.go.jp/shougai/whitepaper/h25hakusho/zenbun/h1_02_00_04.html, (参照 2017-12-05).
- [3] “DAISY”. <http://www.dinf.ne.jp/doc/daisy/about/>, (参照 2017-12-05).
- [4] “ドキュメントトーカー”. <http://www.createsystem.co.jp/main/>, (参照 2017-12-05).
- [5] “インクルーシブ教育システム構築支援データベース”.http://inclusive.nise.go.jp/?page_id=15,(参照 2017-12-5)
- [6] “入学選抜研究機構報告書”. <http://www.dnc.ac.jp/albums/abm.php?f=abm00001112.pdf&n=%E5%85%A5%E5%AD%A6%E8%80%85%E9%81%B8%E6%8A%9C%E7%A0%94%E7%A9%B6%E6%A9%9F%E6%A7%8B%E5%A0%B1%E5%91%8A%E6%9B%B811.pdf>, (2017-12-5)
- [7] Rello, L., & Baeza-Yates, R. : Evaluation of DysWebxia: a reading app designed for people with dyslexia. In Proceedings of the 11th Web for All Conference , 2014 , p.10. ACM
- [8] V. F. Santana, R. Oliveira, L. Almeida, and M. Ito. Firefixia: An accessibility web browser customization toolbar for people with dyslexia. In Proc. W4A '13, Rio de Janeiro, Brazil, 2013.
- [9] L. Rello, M. Pielot, M. C. Marcos, and R. Carlini. Size matters (spacing not): 18 points for a dyslexic-friendly Wikipedia. In Proc. W4A '13, Rio de Janeiro, Brazil,2013.
- [10] L. Rello and R. Baeza-Yates. Good fonts for dyslexia. In Pro

- c. ASSETS'13, Bellevue, Washington, USA, 2013. ACM Press.
- [11] Sampath, H., Sivaswamy, J., & Indurkha, B.: Assistive systems for children with dyslexia and autism. ACM Sigaccess Accessibility and Computing, 2010, (96), 32-36.
- [12] 小山智史. DAISY 風テキストリーダーの開発. 弘前大学教育学部紀要(108), 2012, p.163-169.
- [13] “OrCam”. <https://www.orcam.com/en/>,(参照 2017-12-06)
- [14] “Amazon Polly”. <https://www.polly-jp.net/>,(参照 2017-12-06)
- [15] 廣瀨忍. ディスレクシアについての理解の現状. 岐阜大学教育学部研究報告.人文科学 56(1), 2007, p.205-214.