

## 音声読み上げサービスの機能と精度の比較

山崎 航\*(lz300011@senshu-u.jp),

野口武悟\*(takenori@isc.senshu-u.ac.jp), 植村八潮\*(yashio@isc.senshu-u.ac.jp)

\* 専修大学文学部

### 1. 背景と目的

#### 1.1 研究背景

視覚障害者はそうでない人と同じように墨字（活字印刷された文字）で記されている書籍で読書を行うことは難しい。視覚障害者は読書をする際、墨字の書籍から媒体変換を行った「点字資料」「録音資料」「拡大文字資料」などを利用し、読書を行う。しかしそれらのメディアでは利用できない書籍も少なくない。

そのようなアクセシブルなメディアに対応していない墨字の書籍については、視覚障害者自身が「自炊」し読むことがある。自炊とは紙の書籍をスキャナーなどでOCR（Optical Character Recognition/Reader：光学的文字認識）機能を使用してテキストデータ化し、PCや支援機器などのTTS（Text To Speech：テキスト読み上げ）機能を通して読み上げる方法である。

近年、OCR・TTS機能の技術の発展から2つの機能を一体化させた「文字認識音声読書器」や、スマートフォンのアプリケーションで書籍を撮影し、読み上げる「音声読み上げアプリ（以下、読み上げアプリ）」も登場した。本稿では「文字認識音声読書器」、「読み上げアプリ」を総称して「音声読み上げサービス」とする。

また2000年代から出版業界では電子書籍が台頭し始めた。アメリカのAmazon社のKindleやApple社のiBooks Storeなどの電子書籍は文字の縮小・拡大や音声読み上げなどを利用可能であり、視覚障害者の読書機会拡大に大きく寄与した。また2010年には全国視覚障害者情報提供施設協会と日本点字図書館によって視覚障害者情報総合ネットワークシステム「サピエ」が開始された。本サービスのうち「サピエ図書館」は全国の点字図書館等で作られた点字図書、録音図書に関するデータベースをもとに、貸出し、もしくはダウンロードなどが可能となっている。

#### 1.2 研究目的

前述の通り、近年、電子書籍の台頭などにより視覚障害者の読書環境の向上は進んではいるが、十分ではない側面もある。

電子書籍ストアで販売されている電子書籍の全てが音声読み上げの対象という訳ではな

く、また電子書籍化が実現しない過去の書籍については利用することが難しいことが課題である。

そこで期待されるのが先述した「音声読み上げサービス」である。しかしながら、電子書籍の機能や精度について言及する研究は松原（2017） など散見されるものの、音声読み上げサービス、特に読み上げアプリを扱った研究は存在しない。

そこで本研究では音声読み上げサービスを実際に使用し、読み上げ精度および機能の点からサービスの現状を明らかにする。そのうえで、今日において視覚障害者の読書に最も適しているサービスは何か、サービス使用時の留意点、サービスの今後の課題について考察する。

## 2. 方法

本研究の対象は視覚障害者（全盲、ロービジョン）である。

まず「文字認識音声読書器」、「音声読み上げアプリ」について、機能の面から差異を明らかにする。その後、電子書籍の読み上げ精度を明らかにした松原の研究（2017）を参考に、「文字認識音声読書器」、「音声読み上げアプリ」を使用し、書籍を読み上げた際の誤読率（精度）を調べる。また、YourEyes ボックスを使用した場合の誤読率の変化や標準偏差を求めた。

文字認識音声読書器に関しては神奈川県相模原市立橋本図書館にご協力いただき、株式会社アメディアの「よむべえ」を使用した。「音声読み上げアプリ」は「YourEyes」、「Envision AI」、「Seeing AI」、「サリバン+」、「Microsoft Lens PDF Scanner（以下 Lens）」を使用した（図1）。

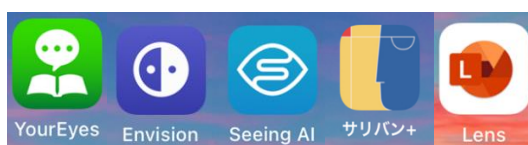


図1 使用アプリケーション一覧

また「読み上げアプリ」の使用端末は「iPod Touch(第7世代)、システムバージョン 14.4.2」を使用し、アプリをAppStore から入手した。

その上で、千葉県立西部図書館に勤務する視覚障害当事者である松井進氏、株式会社ポニーキャニオンが開発した読書支援サービス「YourEyes」の開発担当者である黒澤格氏にヒアリングを実施した。

### 3. 結果

#### 3.1 機能の比較

前述した「よむべえ」と5つの読み上げアプリの特徴の評価・比較を行った(表1)。評価の項目は山口・植村・青木の研究(2012)を参考にした。しかし、山口らの研究は電子書籍を対象としているため、紙の書籍を対象とする本研究では「ネットワーク通信の有無」「画面上のテキストの表示」などを評価項目として加えた。「読み上げ速度の調整」や「声の種類を選択」などは、端末の設定のアクセシビリティ機能によるものではなく、アプリ内での設定が可能かを比較した。

評価は原則○、△、×の三段階で行い、○は「可、もしくは使用にあたって問題がない」、△は「場合により、可」、×は「不可」を意味する。

「基本的な読み上げ」はどれも一定以上の水準は満たしていたが、「Lens」のみ誤読が多く理解不能な文も存在した。

「停止・早送り・巻き戻し」に関しては、「YourEyes」と「よむべえ」のみ3機能全て対応、かつ調節可能だったが、画面上のテキストの表示は対応していなかった(「よむべえ」は専用機器があれば可能)。ロービジョンの方にとってはテキストの表示は重要な要素であることが考えられる。また、ボランティアによる修正の可能性がある「YourEyes」に関しては「抑揚表現の調整」と「間の設定の調整」は現状では×だが、△とした。

各機器やアプリの値段においては「よむべえ」は198,000円とアプリと比べて高価であった。

#### 3.2 精度の比較

まず、(1)各サービスの精度を明らかにするため対象書籍の連続する5ページを読み上げ、1ページあたりの誤読率・誤読数の平均を明らかにした。

次に、(2)読み上げアプリのみに視覚障害者でも書籍を正面から撮影することが可能になるYourEyesボックス(図2)を使用し、不使用の場合との1ページあたりの誤読数・誤読率を比較する。

アプリの場合、書籍を撮影する際に手ブレなどが生じ、それがOCR技術に影響を及ぼすことが考えられる。そこで、(3)対象箇所の中から1ページに限定し、同ページでボックス使用時・未使用時の2パターンで読み上げを各10回行い、誤読数とその標準偏差値を求めた。ここから撮影方法がどれだけ誤読数に影響を及ぼすか、またどのアプリが最も安定した精度のパフォーマンスかが明らかになる。

表1 音声読み上げサービスの機能の比較

	Your Eyes	Envision AI	Seeing AI	サリバンプ	Microsoft Lens	よむべえ
使用可能端末	iOS	iOS/Android	iOS	iOS/Android	iOS/Android	
音声エンジン	Read Speaker	VoiceOver/Talkback	Voice Over	VoiceOver/Talkback	イマーシブリーダー	
月額料金	550円	200円	無料	無料	無料	
ネットワーク通信	必要	必要	必要	必要	必要	不要
基本的な読み上げ	○	○	△	○	×	○
漢字の読み	○	△	△	○	×	○
感嘆符/疑問符の読み上げ	○	○	△	○	○	○
括弧の読み上げ	○	○	○	○	○	○
画面上のテキストの表示	×	○	○	○	○	△
テキストの拡大・縮小		○	○	×	○	△
「/」 /…の読み	○	○	○	○	○	○
英語読み	○	○	×	△	×	○
声の種類を選択	○	○	○	×	○	×
改行/字下げ	×	△	×	△	○	○
停止・早送り・巻き戻し	○	△ (停止のみ)	△ (停止のみ)	×	△ (停止のみ)	○
読み上げ速度の調整	○	○	×	×	○	○ (10段階)
抑揚表現の調整	△ (ボラ ン ティア)	×	×	×	×	×
間の設定の調整	△ (ボラ ン ティア)	×	×	×	×	×



図2 YourEyes ボックス

文庫本・単行本の1ページあたりの誤読率・誤読数は表2の通りである。( )内は誤読数である。

サービスごとに誤読率を見ると、「YourEyes」、「サリバンプ」は文庫本、単行本ともに誤読数はおよそ1.0~1.5%と少ないのに対し、「SeeingAI」、「Lens」は比較的多い結果になった。有料サービスである「EnvisionAI」は文庫本では2番目に誤読が多いという結果となった。

誤読において全サービス内で共通したものとして「ページ移りの前後の一行は誤読する可能性が高い」点が見られ、誤読例としては『つ』を『0』、『摺鉢(すりばち)』を『しゅうばち』と読むような間違いが多く、特に「Envision AI」、「Seeing AI」、「Lens」に見られた。

表2 文庫本、単行本の1ページあたりの誤読数

	YourEyes	Envision AI	Seeing AI	サリバンプ	Lens	よむべえ
文庫本	0.8% (3.9)	3.2% (15.4)	3.1% (14.7)	1.0% (4.9)	4.5% (21.9)	1.8% (8.9)
単行本	1.2% (5.7)	2.2% (10.3)	2.7% (12.9)	1.5% (7.3)	3.1% (14.8)	2.3% (10.9)

次に YourEyes ボックスの使用時と不使用時の文庫本の誤読数の比較を行った (表 3)。「YourEyes」に関してはボックスありの場合の方が誤読数が増えたが、他 2 つのアプリは減少した。「YourEyes」の誤読が増加した理由としてフラッシュの反射が挙げられる。

YourEyes ボックスは底にアクリル板が設置されており、端末で紙面を撮影する際にカメラのフラッシュが反射し、紙面の一部が白飛びしてしまうことがある (図 3)。そのためその部分の文字の OCR がうまく作動しなかったことから、誤読数が増加してしまったと考えられた。

表 3 文庫本の YourEyes ボックス無しとありの誤読率・誤読数の比較

	YourEyes	Envision AI	Seeing AI	サリバンプ	Lens
ボックス無し	0.8% (3.9)	3.2% (15.4)	3.1% (14.7)	1.0% (4.9)	4.5% (21.9)
ボックスあり	0.9% (4.5)		2.7% (13.1)		4.1% (19.9)

※アプリの仕様上ボックスを使用できなかったものは空欄になっている。



図 3 YourEyes ボックス使用時の反射

YourEyes ボックス不使用時、使用時の標準偏差値は表 4 のとおりである。

標準偏差とは分散の正の平方根で、データの分布の散らばり (全体の平均からどの程度ばらつきがあるか) の程度を表す。つまり標準偏差値が大きければデータにばらつきがあり、小さければばらつきの具合が小さいということになる。

「よむべえ」は標準偏差値 0.8 で常に安定した読み上げを行うことが可能であり、次いで

「サリバンプ」、「EnvisionAI」、「YourEyes」の順に誤読数のばらつきは大きくなった。しかしボックスを使用したことにより端末と書籍が固定され誤読数のばらつきが抑えられた。その結果標準偏差値は減少した。ここから YourEyes ボックスは精度に大きく影響をもたらすものという以上に安定性を向上させるものであることが明らかになった。

表 4 各サービスの標準偏差値

	YourEyes	EnvisionAI	SeeingAI	サリバンプ	Lens	よむべえ
ボックス無し	3.5	3.1	6.7	2.5	10	0.8
ボックスあり	0.5		2.1		2.3	

※アプリの仕様上ボックスを使用できなかったものは空欄になっている。

### 3.3 ヒアリング

#### (1) 黒澤格氏へのヒアリング

「YourEyes」の開発担当者である株式会社ポニーキャニオンの黒澤格氏にヒアリングを行った。

黒澤氏は「YourEyes」の強みについて「今そこにある書籍をすぐに読める」という点と、「ボランティア」を挙げられた。

点字資料や DAISY などの音訳資料といった、既に媒体変換された書籍のみ読むことが出来る、もしくは読みたい書籍があっても媒体変換を待つよりほかはないという受動的な読書もある。しかし「YourEyes」は自らが読みたいものを読みたいときに読むことが出来る即時的・能動的かつ主体的なメディアである。

また「YourEyes」はボランティアによる読み上げの修正が存在する。現在の OCR、TTS 技術では少なからず読み間違いは存在し、視覚障害者の読書を阻害しているが、修正をし、正しく読まれることが可能になれば「YourEyes」はさらに高精度になり、より良いサービスになる。

#### (2) 松井進氏へのヒアリング

千葉県立西部図書館勤務で自身も視覚障害者である松井進氏にもヒアリングを行った。

音声読み上げサービスでは上記の黒澤氏へのヒアリングにもあった「即時性」を考慮した上でサービスの「使い分け」が求められる。

使い分けとは例えば書籍を読む場合には裁断による自炊を使用し、手軽に A4 サイズの書類などを読みたい時や自炊に必要な機器がそろっていない場合には読み上げアプリを使用するなどである。このように視覚障害者の読書方法の選択肢を増やすことが求められる。

また特にアプリに関しては「撮影の行いやすさ」も求められる。視覚障害者が書籍の紙面をスマートフォンで撮影する場合、ブレや紙面の見切れ、上下左右の反転など OCR 機能に

支障をきたす可能性が高くなる。そこで YourEyes ボックスなどの補助機器を使用することが好ましい。

#### 4. 考察と結論

本稿では、ここまで「音声読み上げサービス」の特徴や精度の比較、音声読書アプリ「YourEyes」の開発に携わった方にお話を伺うことなどを通し、読み上げサービスの現状を検討してきた。

これらを通して3サービスの長所と短所を表5にまとめた。

表5 読書における読み上げサービスの長所と短所

読み上げサービス	長所	短所
自炊	<ul style="list-style-type: none"> <li>・ページめくりが不必要</li> <li>・機能が充実している（汎用性が高い）</li> </ul>	<ul style="list-style-type: none"> <li>・比較的高価</li> <li>・持ち運び不可</li> </ul>
文字認識音声読書器	<ul style="list-style-type: none"> <li>・機能が充実している（汎用性が高い）</li> </ul>	<ul style="list-style-type: none"> <li>・高価</li> <li>・持ち運び不可</li> <li>・ページめくりが必要</li> </ul>
音声読み上げアプリ	<ul style="list-style-type: none"> <li>・安価</li> <li>・持ち運び可能</li> <li>・入手方法が簡単</li> <li>・即時性</li> </ul>	<ul style="list-style-type: none"> <li>・精度がまちまち</li> <li>・ボックスが必要</li> <li>・ページめくりが必要</li> </ul>

上記の長所短所を踏まえ、以下のような3つが明らかとなった。

- (1) 墨字の読書において、精度の観点では理想のアクセシブルな読み上げサービスは「音声読み上げアプリケーション」である。
- (2) 読み上げアプリを使用する際は YourEyes ボックスのような OCR 機能をサポートする機器が必要不可欠である。
- (3) 読み上げサービスは一長一短であり「使い分け」が求められる。

以上3つの詳細を以下で述べる。

##### 4.1 理想の読み上げサービス

安価、持ち運びが可能、入手手段が比較的容易という点から現状では読み上げアプリが最もアクセシブルであるという結論に至った。また精度においても読み上げアプリは誤読率が少なく同様に最も適していると言える。中でも特に「YourEyes」、「サリバソ+」は誤読数の少なさ、YourEyes ボックスなしで読み上げを行った際の誤読数のばらつきから



読書に適していると言える。

#### 4.2 YourEyes ボックスの必要性

視覚障害者にとってアプリで書籍を撮影した際、対象箇所を撮影できているのか確認することは難しい。そのため現状では YourEyes ボックスのような画角や位置・距離を固定できる機器が必要となる。

しかし YourEyes ボックスを使用することによって標準偏差の値は改善されたが、誤読数は全てにおいて同様とはならなかった。OCR で読み取る際にカメラのフラッシュによって文字が白飛びしてしまった為である。改善策としては側面の材質を透明なアクリルにすることで外部からの光を取り込むことが考えられる。

#### 4.3 使い分け

各サービスに短所が存在し、完璧な音声読み上げサービスは今日には存在しないため、必要とされるのは「使い分け」である。特に視覚障害者一人ひとりのニーズに合わせた「使い分け」、読み上げを行う際の状況や目的のニーズに合わせた「使い分け」の2つである。

個人のニーズに合わせた「使い分け」について、項目は例示に過ぎないが、重要視する項目ごとの最適なサービスを表6にて示した。

「費用」や「持ち運び」を重要視した場合アプリは最適なサービスだが、「安定した精度」という点では自炊や音声読書器が最適である。「手軽さ」という観点においてもアプリはすぐに使えて手軽ではあるが、読み上げの際にページめくりを手動で行わなければならない。その点自炊は、使用する機器にもよるが裁断さえ行えば自動でページ送りやOCR、TTSを行うものもあるため、異なる意味で手軽である。

また、アプリの中でもテキストを画面上に表示する機能や聞き直しに有効な読み上げ音声の巻き戻しや早送りなどの有無によって個々人が最も重要視する要素によって使い分けが求められる。

表6 ニーズに合わせた使い分け

重要視する項目	自炊	文字認識音声 読書器	音声読み上げ アプリ
費用			最適
持ち運び			最適
精度			最適
安定した精度	最適	最適	
汎用性	最適 (読書時)		最適

次に、状況や目的に合わせた「使い分け」には、読み上げ対象の重要度が大きく関係する。

視覚障害でない人は読むものによって初めから終わりまで精読する場合や読み捨てる場合など、読む労力の使い分けを普段から行っている。これは視覚障害者においても同じである。書類などを早読したい場合には少しくらいの誤読は許容し、ボックスを使用せずにアプリで読み上げることもあるだろう。反対に詳説などを精読したい場合は時間をかけてもボックスを使用する、もしくはより良いソフトを使用した自炊を行うなど工夫をすることも可能である。

「精度」や「即時性」など一つの観点に縛られるのではなく、各々の異なる多様なニーズにこたえられるよう読み上げサービスを提供することが今日の理想のアクセシブルな読書メディアへの第一歩であると言える。

#### 【参考文献】

- ・高橋陽一郎ほか『詳説数学 B』新興出版社啓林館、2012
- ・野口武悟、成松一郎編著『多様性と出会う学校図書館：一人ひとりの自立を支える合理的配慮へのアプローチ』読書工房、2015
- ・野口武悟、植村八潮編著『改訂 図書館のアクセシビリティ：「合理的配慮」の提供へ向けて』樹村房、2021
- ・松原聡『電子書籍アクセシビリティの研究：視覚障害者等への対応からユニバーサルデザインへ』東洋大学出版社、2017
- ・山口翔、植村要、青木千帆子「視覚障害者向け音声読み上げ機能の評価：電子書籍の普及を見据えて」『情報通信学会誌』30 (2)、2012

#### 【謝辞】

ご多忙のところ、調査にご協力いただいたみなさまに、心より御礼を申し上げます。