

テキストを指でなぞる速度と音声の速さが連動する電子書籍読み上げシステムの開発と評価

松本 愛^{1,a)} 高田 峻介¹

概要：オーディオブックのような音読システムを用いると本を音声で楽しむことができる。しかし、音声の速度を変更するには再生速度を都度クリックして切り替える必要があり、内容に沿って逐次的に読む速さを変化させるのは難しい。そのため、黙読時のように文中の人物や情景・場面によって読者の好みに合わせて緩急をつけた読み方ができない。そこで本研究では、音読システムでも読者の好みを反映させた読みができるよう、画面上の文字をなぞる速さに比例した速度の音声再生が再生される電子書籍リーダアプリを提案する。さらに、文字だけの条件と音声の速さをボタンで変更できる条件と比較し、読了時間や文章の理解度、使用感にどのような差があるか検証する。

1. はじめに

文部科学省は『子どもの読書活動は、言葉を学び、感性を磨き、表現力を高め、創造力を豊かなものにし、人生をより深く生きる力を身に付けていく上で欠くことのできないもの』として、読書を推奨している [1]。読書には黙読や音読、読み聞かせなど様々な方法がある。さらに電子端末が日常的に用いられるようになり、電子書籍やオーディオブックなど、新たな読書の手法が登場した。

オーディオブックのような音読システムを用いると本を音声で楽しむことができる。しかし、音声の速度を変更するには再生速度を都度クリックして切り替える必要があり、内容に沿って逐次的に読む速さを変化させるのは難しい。そのため、黙読時のように文中の人物や情景・場面によって読者の好みに合わせて緩急をつけた読み方ができない。

そこで本研究では、音読システムでも読者の好みを反映させた読みができるよう、画面上の文字をなぞる速さに比例した速度の音声再生が再生される電子書籍リーダアプリを開発した。アプリ画面を図1に示す。提案手法の有効性を示すため、文字だけの状態と音声の速さをボタンで変更できる状態と比較し、読了時間や文章の理解度、使用感にどのような差があるか検証する。

2. 関連研究・手法

提案手法は電子書籍における音読体験を拡張するものである。そのため関連研究として、電子書籍の拡張技術お



図1: アプリケーションの動作イメージ

び電子書籍における音読に関する手法を挙げ、本研究との差異を述べる。

2.1 電子端末を使用する読み上げツール

近年は本をアナウンサーや声優が音読した音声配信するオーディオブックのサービスも利用されている (audible^{*1}など)。また、電子端末で使用できる読み上げツールなどもある [2][3]。

これらの手法では音声を都度クリックで切り替える必要がある。提案手法は、内容に沿って逐次的に速度を変化できる点で異なる。

¹ 神戸市立工業高等専門学校 電子工学科

^{a)} r120337@g.kobe-kosen.ac.jp

^{*1} Amazon, <https://www.audible.co.jp/>, (Accessed: 2024/12/24)

2.2 電子書籍を拡張する研究

電子書籍の特性を活用した研究として、ピンチイン/アウトにより文章の拡大・縮小だけでなく、文章の内容を要約度合いを変化させる Text Level Of Detail[4] や、読者が文章に対して投稿したコメントの量などを可視化させることで、文章中の行間をメディアとして活用させるもの [5] がある。また、文章の見た目へアプローチしたものとして、ページごとに出てくる単語から連想させる色を背景色として表示するシステム [6] や、電子書籍に図書の推薦システムを搭載する研究 [7] などがある。

これらの研究に対し、提案手法は電子書籍の音読へアプローチする点で異なる。

2.3 電子書籍の音声変化やシステムに関する研究

電子書籍の音声に着目した研究として、文章の内容を分析し合成音声用の制御メタデータを生成することによる自然で聞きやすい読み上げ機能 [8] や、電子コミックや挿絵のある小説のキャラクターの画像からキャラクターごとに適した声優を分類するシステム [9] がある。

これらの研究は電子書籍に沿った音声を生成するものであり、提案手法は読者の好みに応じて音声の再生速度を変えるものであるためアプローチが異なる。また、提案手法はこれらの手法と組み合わせて使用できる手法である。

3. 提案手法

タッチパネル端末上で、文章を指でなぞる速度に比例して読み上げ速度が変化する電子書籍リーダアプリケーションを実装した。

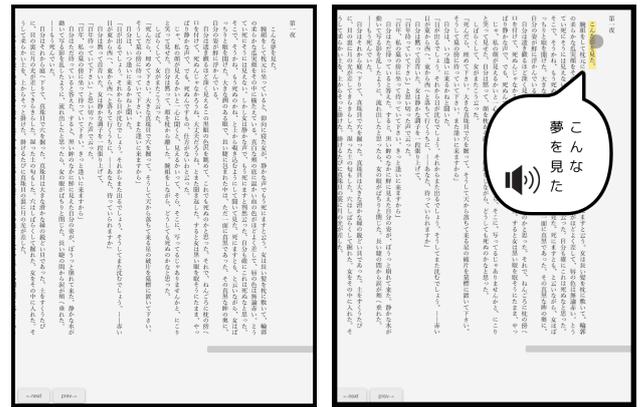
3.1 設計

図 2(a) に示すように、操作を何もしていない状態では、文章とページ移動ボタンのみが表示される。読者が画面に表示された文字をなぞると、図 2(b) に示すようにタッチした箇所の文章がハイライトされ、文章に対応した読み上げ音声が再生される。アプリケーションのフローチャートを図 3 に示す。音声の速度は画面を上から下へなぞる速さに比例するため、ゆっくりなぞると音声もゆっくりに、速くなぞると音声も速くなる。

なぞる指と読んでいるテキストが重ならないように、指を動かしても再生される箇所が変化しないように、再生する音声は画面から指を離すまで文章の流れに沿う。

3.2 実装

テキストの処理には Python を、電子書籍リーダは PC やスマートフォンといった異なる端末、および OS 上で動作するように、HTML/CSS/JavaScript を用いて Web アプリケーションとして実装した。また、読み上げ音声の生



(a) 操作をしていない場合 (b) 文字をなぞっている場合

図 2: アプリケーションの操作での変化

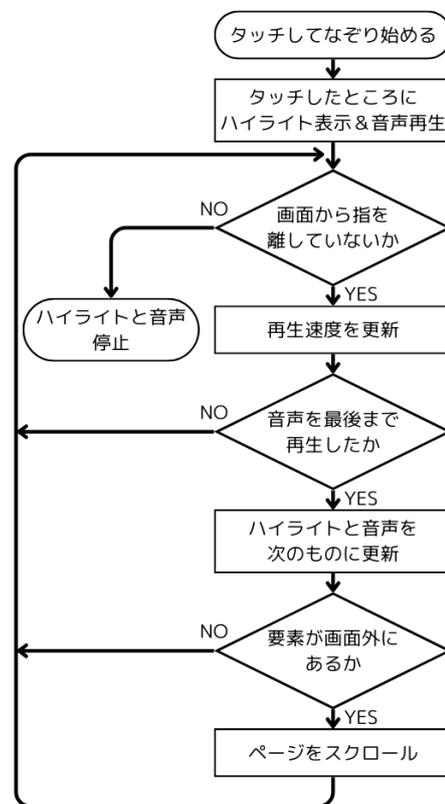


図 3: 動作のフローチャート

成には合成音声ソフトである VOICEVOX *2を使用した。用いた合成音声モデルはずんだもんである。

Python 上のテキスト処理の流れを図 4 に示す。まず、青空文庫 *3からダウンロードした作品の HTML データから本文テキストのみを抽出し、文章を読点、句読点、改行、全角スペース、鍵括弧ごとに区切った文章を配列に格納した。それから、電子書籍リーダ用の HTML ファイルおよび音声生成に使用するためのテキストファイルを作成した。その後、各文章の読み上げ音声ファイルを生成した。

*2 Hiroshiba Kazuyuki, <https://voicevox.hiroshiba.jp/>, (Accessed: 2024/12/24)

*3 青空文庫, <https://www.aozora.gr.jp/>, (Accessed: 2024/12/24)

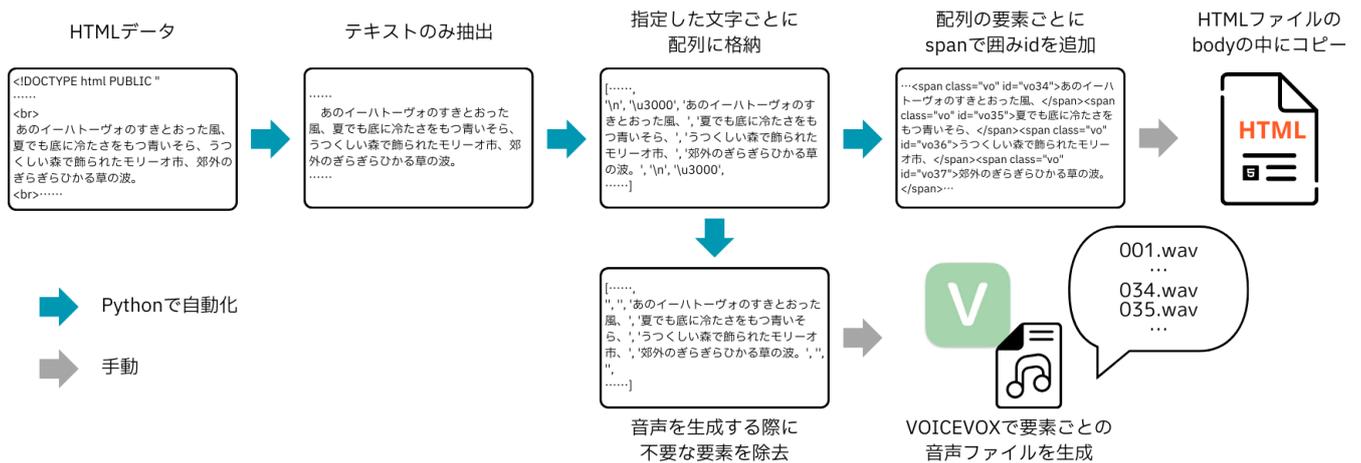


図 4: テキストの処理の流れ

3.3 なぞる動きと再生速度連動

なぞる動きから再生速度 v_p を算出する方法を示す。

まず、再生速度の上限 S_{MAX} と下限 S_{MIN} 、スケーリング係数 k の値を決定しておく。タッチ移動の動き (JavaScript の touchmove イベント) が発生すると、タッチしている箇所の Y 座標とミリ単位の時間を取得する。画面に指が触れている間は座標と時間を取得し続け、前回との差分を計算する。座標と時間の差分を $\Delta Y, \Delta t$ と表すとそれぞれは、

$$\Delta Y = Y(t) - Y(t - \Delta t)$$

$$\Delta t = \text{取得した時刻} - \text{更新前の時刻}$$

で求められる。

垂直方向の移動速度 v_y は、

$$v_y = \frac{\Delta Y}{\Delta t} [\text{pixel/ms}]$$

で算出できる。上から下への指の動きのみを音声の再生速度に反映させるために、以下の速度計算は $\Delta Y > 0$ となる時のみ行う。 v_y の値をそのまま使用すると再生速度があまり変化せず、た変化がきこなくなるため、スケーリングとスムージングをする。スケーリングは v_y にスケーリング係数 k を掛けることで行う。 k はデバイスや文字のサイズに合わせて手動で値を変える。

求めた値を $v_{scaling}$ とし、スムージングした速度 (v_{smooth} とする) を以下の式で求める。

$$v_{smooth} = \frac{v_{scaling} + v_p}{2}$$

v_{smooth} と S_{MAX} , S_{MIN} を用いると v_p は、

$$v_p = \min(S_{MAX}, \max(S_{MIN}, v_{smooth}))$$

で求められ、読み上げ音声の再生速度として値を更新する。

4. 実験

実装した電子書籍リーダーを用いて、実験協力者による比較実験を行った。

4.1 調査環境および実験協力者

実験に音声を用いるため、調査は静かな部屋で行った。参加者には、パソコンに接続したタッチパネルディスプレイ (InnoView, INVPM204, 39.3×7.4×26.6 cm, 3840×2160 Pixels) を使用してもらった。ディスプレイは縦向きに設置し、Chrome のデベロッパーツールでタブレット (13 インチ) ほどのサイズのアプリケーション画面を表示した。

調査は神戸高専電子工学科の 4・5 年生 6 名 (男性 5 名、女性 1 名、19-20 歳) を対象に実施した。調査で使用した作品を読んだ経験がある参加者はいなかった。また、全員が電子書籍を使用した経験があった。

4.2 調査方法

調査協力者に小説 3 作、評論 3 作を、以下の 3 つの条件のいずれかで読んでもらった。

パターン 1 音声なし

パターン 2 音声あり・ボタンで再生速度変更

パターン 3 音声あり・なぞる速さで再生速度変更

作品とパターンの組み合わせや作品を読む順番は、偏りがないよう参加者ごとに行列で順番に変化させた。

1 つの作品を読み終わるごとに、読み終わるまでの時間を記録し、作品への興味関心とユーザビリティに関するアンケートに回答してもらった。アンケートの内容は後述する。パターンごとの結果を比較し、アプリケーションがユーザーに与える影響を調査する。

4.3 使用した文章

青空文庫に HTML データがある文章を使用した。また、読む作品の内容が似たようなものになるように、小説 3 作と評論 3 作ごとに同一の著者のコンセプトや内容の似た 2000-4000 文字の作品を選んだ。表 1 に文章の一覧を示す。

表 1: 調査で使った文章

番号	タイトル	抜粋箇所	作者	文字数
小説 1		第一夜・第二夜		3823
小説 2	夢十夜	第三夜・第四夜	夏目漱石	3350
小説 3		第七夜・第十夜		3568
評論 1	科学者とあたま	-		3869
評論 2	数学と語学	-	寺田寅彦	3584
評論 3	知と疑い	-		2135

アンケート内容	
Q1	文章の理解関心について
Q1-1	この小説・評論は分かりやすかったですか
Q1-2	この小説・評論は面白い・興味深かったですか
Q1-3	この小説・評論の内容を理解できましたか
Q1-4	この小説・評論は問題なく読めましたか
Q1-5	分かりやすさや面白さに関して理由などがあれば記載してください
Q2	アプリケーションの使用感について
Q2-1	このシステムを頻繁に使いたいと思う
Q2-2	このシステムは不必要に複雑だと思った
Q2-3	このシステムは使いやすいと思った
Q2-4	このシステムを使うのに技術専門家のサポートが必要とするかもしれない
Q2-5	このシステムにあるさまざまな機能がよくまとまっていると感じた
Q2-6	このシステムでは、一貫性のないところが多かったとおもった
Q2-7	たいいていのユーザは、このシステムの仕様方法について、素早く学べるだろう
Q2-8	このシステムはとても扱いにくいと思った
Q2-9	このシステムを使うのに自信があると感じた
Q2-10	このシステムを使い始める前に多くのことを学ぶ必要があった
Q2-11	アプリケーションの使用に関して何か意見があれば記入してください

図 5: アンケートの内容

4.4 質問項目

アンケート内容を図 5 に示す。質問 1 で文章内容に対する参加者の興味関心を、質問 2 でアプリの使用感を回答してもらった。質問 2-1 から 2-10 まではシステムユーザビリティスケール (SUS)[10] の規格を基にした。

5. 結果と考察

4 章にて実施した実験の結果と考察を述べる。

5.1 読む時間に与える影響

得られた値から 1000 文字あたりの読む速度を計算し条件ごとにまとめた。結果を表 2 および図 6 に示す。

これらより、文章のみ条件が最も読む速度が早く、音声機能を用いる場合が遅く、提案手法も音声のみ条件も速度にあまり差が無いことが分かる。分散値は提案手法が音声あり条件よりも大きい。これは、提案手法で文章を読む際に文章のすべてをなぞる参加者と一部だけをなぞる参加者がおり、読む速度に差が生まれたためと考えられる。

表 2: 読む速度の平均値と標準分散

ジャンル	パターン	平均値	標準分散
小説	文章のみ	01:51	1061.81
	音声あり	02:40	1264.92
	提案手法	02:41	1940.25
評論	文章のみ	02:15	1717.14
	音声あり	02:37	347.67
	提案手法	02:32	1191.22

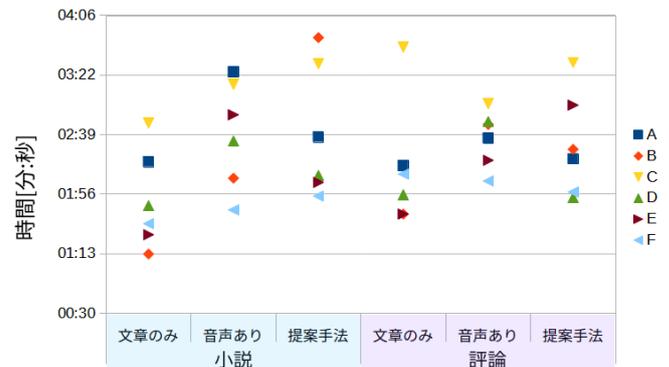


図 6: 読む速さの変化

5.2 文章の理解関心に与える影響

Q1-1 から Q1-4 の平均値を条件ごとに求めた。結果を図 7 および図 8 に示す。

結果より、理解関心に関してはパターンによって大きな差は生まれていないことが分かる。音声機能を用いると関

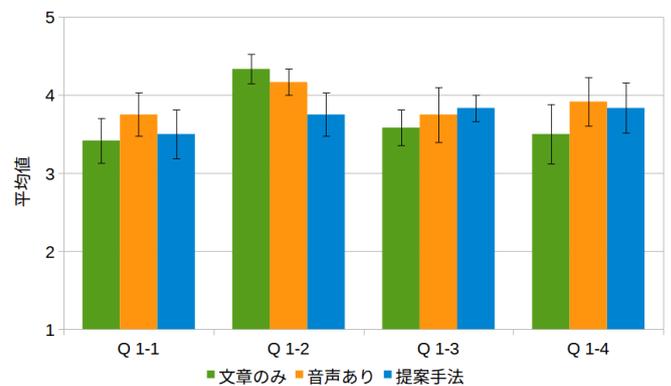


図 7: 興味関心の平均値 (全体)

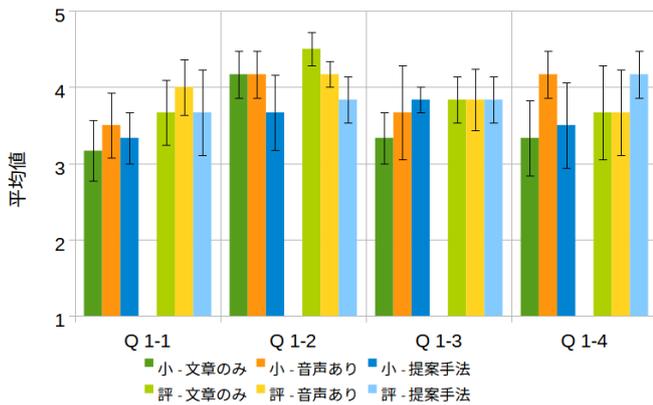


図 8: 興味関心の平均値 (ジャンルごと)

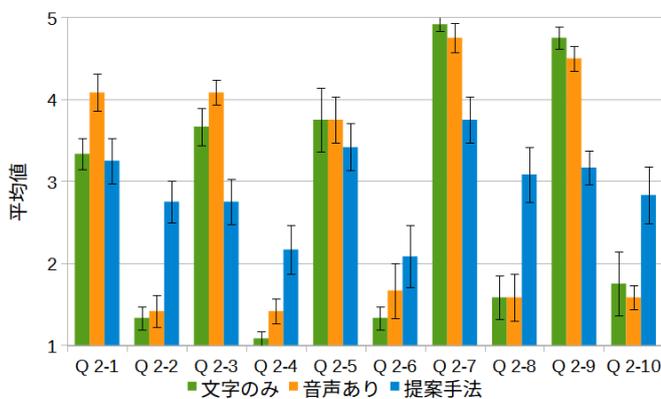


図 9: 使用感に対する回答の平均値

心度は減少するが理解度は増加する傾向が見られる。自由記述より、音声機能がない場合は漢字の読みがわからないという意見もあった。また、図 6 と比較すると、読む速度が遅くなった分、理解度が上昇した可能性が示唆された。

5.3 アプリケーションの操作性

Q2-1 から Q2-10 の平均値を図 9 に示す。結果から算出した SUS の値は文章のみと音声ありが 83 で提案手法が 58 である。提案手法は操作方法が独特で、慣れるまでに時間を要するため値が低くなったと考えられる。そのことを述べる意見もいくつかあった。今後は、操作に慣れる時間を設ける、操作性の向上が必要だと考えられる。

6. 議論

実験に用いる作品および提案手法の利用先を議論する。

6.1 作品の選定

実験協力者は使用した作品を難解と感じている事が多かった。そのため、文章の意味を理解することに集中してしまい、提案手法を使用して作品を楽しむ事がなされなかった可能性がある。今後は理解がしやすい児童文学やライトノベルなどを用いて実験を行う。

6.2 ディスレクシアとそれについての研究

提案手法は、文字を読む黙読体験と音声聞く音読体験を組み合わせた手法である。また、既存の音読手法では難しかった黙読における読者の好みに合わせた読む速度やテンポを取り入れることができる手法である。さらに、実験時の意見にもあった通り、例えば読みが難しい感じなども音で聞くことができる。

調査によると、読書に苦手意識を持つ人も一定数おり、その理由として、文字を読むのが苦手であることや文章を読もうとすると疲れることが挙げられている [11]。そういった人は識字障がい (ディスレクシア) という学習障がいの一つに含まれる場合もある。ディスレクシアは文字の認識や音韻処理が苦手という特性を持ち、人口の数%いるとされる [12]。また、読みに困難がない児童とある児童でグループを分け、紙の教科書での音読とデジタル教科書での音読と音声読み上げで目の動きを比較した研究では、デジタル教科書を用いると 2 つのグループでの読み速度の差が少なくなるということが報告されている [13]。

本研究はこのように読みに困難を持っており、読書が苦手であったり、敬遠しているような読者でも本を好みの速度で読むことができる可能性がある。

7. まとめ

音声を用いて読者の好みを反映させた読みを行える手法として、画面をなぞる速度と読み上げ音声の再生速度が連動する web アプリケーションを作製した。アプリケーションの効果を確認するために、文字のみ、音声あり、提案手法の 3 つの条件で小説と評論を 6 人に読んでもらう比較実験を行った。その結果、同じ音声を用いる場合でも再生速度ボタンで切り替えるよりなぞって変化させる方が読む時間が短くなった。また、どの手法でも作品への理解関心は大きく変化しないことが分かった。

しかしながら、SUS の値は提案手法が他の 2 条件より低い結果となった。改善案もいくつか寄せられた。操作方法に慣れるまでに時間がかかる事や、使用した作品が実験協力者にとって難しかったことが原因であると考えられる。今後はアプリケーションの操作性を改善し、分かりやすい作品を使用した実験を行う。

参考文献

- [1] 文部科学省. 子どもの読書活動の推進に関する基本的な計画, 2023.
- [2] 日本 DAISY コンソーシアム. Daisy. <https://www.japandaisy.org/>(Accessed: 2024/12/24).
- [3] Google. Talkback. <https://www.youtube.com/watch?v=MEp9DtiTw3s>(Accessed: 2024/12/24).
- [4] 武田港, 望月茂徳, 川村健一郎. Text Level of Detail: ズーム動作と階層構造による新しい読書体験の為の 3D 文章システム及びコンテンツの提案. 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2012, No. 8, pp.

1–7, 10 2012.

- [5] 中村将達, 西田健志. 行間を拡張する能動的な読書の提案. エンタテインメントコンピューティングシンポジウム 2013 論文集, Vol. 2013, pp. 146–152, 09 2013.
- [6] Takuto Kamiura, Shohei Yamada, and Naiwala P. Chandrasiri. A System for Generating Background Colors for e-Books, Based on its Text. In *2019 International Conference on Computing, Networking and Communications (ICNC)*, pp. 209–213, 2019.
- [7] Chifumi Nishioka, Sanae Fujita, Takashi Hattori, Tessei Kobayashi, Futoshi Naya, and Hiroaki Ogata. A Picture-Book Recommender System for Extensive Reading on an E-Book System. pp. 690–692. Society for Learning Analytics Research (SoLAR), 3 2020.
- [8] 布目光生, 鈴木優, 森田真弘. 自然で聞きやすい電子書籍読み上げのための文章構造解析技術. 東芝レビュー, Vol. 66, No. 9, 2011.
- [9] 齊藤彰吾, 大井翔, 佐野睦夫. 2次元キャラクタにおける音声生成モデルの検討. 2022年度 情報処理学会関西支部支部大会 講演論文集, Vol. 2022, p. 4, 09 2022.
- [10] John Brooke. SUS: A 'Quick' and 'Dirty' Usability Scale. In Patrick W. Jordan, Bruce Thomas, Bernard A. Weerdmeester, and Ian Lyall McClelland, editors, *Usability Evaluation in Industry*, chapter 21, pp. 189–194. Taylor and Francis, June 1996.
- [11] 日本財団. 18歳意識調査「第30回 – 読む・書く –」, 2020. https://www.nippon-foundation.or.jp/wp-content/uploads/2020/10/wha_pro_eig_152.pdf (Accessed: 2024/12/24).
- [12] Akira Uno, Taeko N Wydell, Noriko Haruhara, Masato Kaneko, and Naoko Shinya. Relationship between reading/writing skills and cognitive abilities among Japanese primary-school children: Normal readers versus poor readers (dyslexics). *Reading and writing*, Vol. 22, pp. 755–789, 2009.
- [13] 氏間和仁, 今津麻衣. 紙教科書とデジタル教科書の読みの際の眼球運動の比較: 読みの困難度に着目した予備的検討. 広島大学特別支援教育実践センター研究紀要, No. 22, pp. 1–8, 03 2024.