

英語多読能力推定のための適応型試験システム開発

竹田 響¹ 新村 正明²

概要：英語教育の一環として特に日本の教育機関を中心に英語多読学習法の一つである SSS 式多読法が広まっており、この学習法は SSS が掲げる多読 3 原則に従い英語多読本を大量に読むというものである。この学習法の問題点は自身の多読能力を定量的に把握する方法が存在せず、初学者は自身のレベルの把握のために余分な読書を行う必要があることである。そこで本研究室の足立らによる研究では英語多読学習者の多読レベル把握のための適応型試験を実施するシステムである ERFPT(Extensive Reading Foundation Placement Test) が提案された。このシステムは項目応答理論に則った適応型試験を実施することで受験者の英語多読レベルを推定・提示し、受験者は提示されたレベルをもとに多読を行うことで英語多読学習における負担を軽減させることを目的としている。このシステムは Web アプリケーションとして提供しており、試験の実施や成績の管理などの機能がある。先行研究ではシステム利用者の情報を全て一つのサーバで管理しており、複数の組織で利用することの考慮がなされていない。さらに先行研究では試験結果が英語多読学習の支援に繋がらないという問題があり、この原因の一つに本学で利用している英語多読学習管理システムとの連携が取れないことが挙げられた。そこで本研究では先行研究で提案された適応型試験を実施するアプリケーションを複数の組織で利用することを考慮した構成で再設計・開発を行った。さらに適応型試験の結果から推定された受験者のレベルを英語多読学習支援に繋げるための手法の提案とその評価を行った。

キーワード：英語多読学習, Web アプリケーション, docker, 適応型試験, 回帰分析

1. はじめに

英語能力の向上に有効な学習法である多読学習を英語教育の一環として取り入れる場合、学習効果を得るためには学習者に多読を継続させる工夫が必要となる。

特に日本の教育機関を中心に広まっている SSS 式多読法 (Start with Simple Stories 式多読法) では SSS が掲げる多読 3 原則に従い「適切な本 (適切なレベルでかつ趣味趣向に合う本)」を多量に読むことを目的としている。この学習法における適切な本を借りるために学習者自身が適切なレベルを正しく認知した上で趣味趣向に合った本を探せる環境を用意する必要がある。そこで、本学の SSS 式多読法を取り入れた講義に、英語多読学習の支援を目的としたシステムである ERS(Extensive Reading System) の導入を行った。これにより趣味趣向に合った本を探せる環境を用意することができ、多読に対するモチベーションの維持・向上等に一定の効果を得ることができた。

また、SSS 式多読法において学習者は自身の適切なレベルを把握するために難易度の低い多読本から読み始め徐々に難易度を上げていく必要がある。そのため特に英語多読の初学者は自身の適したレベルの把握のための読書を行う必要があり、多読学習を開始するまでの時間的・労力的に負担がかかることが課題となる。そこで足立らはこの課題に対して学習者に適切なレベルを認知させることを目的とした多読のレベルを測定するためのテストである ERF Placement Test(以降は ERFPT と記述する) を提案し、Web アプリケーションとして開発・提供した [1]。しかし、この ERFPT には複数組織での運用を考慮した設計がなされていないことや、学習者のレベルの推定結果を多読学習に活用されていないという問題がある。

本報告ではこの問題点を解決するために ERFPT の再設計と、ERFPT で推定された学習者の能力を多読学習に活用するために実装した機能について記述する。さらに ERFPT で推定されたレベルを学習者が多読本の選定時に活用するためには推定値の変換が必要となり、その変換手法と変換後の値について評価を行った。

¹ 信州大学大学院 総合理工学研究科
Graduate School of Science and Technology, Shinshu University

² 信州大学
Shinshu University

2. SSS 式多読法

SSS 式多読法とは SSS が提唱する絵本などの非常に易しい本から多読をはじめ、100 万語単位で多読を行うことを目標とした多読法であり、ERF(Extensive Reading Foundation) が定める Graded readers[2] のようにレベル分けされた本を教材として利用し、易しい本から難しい本へと徐々にステップアップすることによって読解力や語学力を養うことを目標としている。SSS では SSS 式多読法による学習効果の理由を「楽しいからあきない」、「自分のペース・好みによって本が読める」、「今までの英語学習に比べて、10 倍から 1000 倍の英文に触れることができる」としている [3]。さらに多読を長く続けるためには「楽しくスラスラ読む」必要があることから「原則 1：辞書は引かない」、「原則 2：わからないところは飛ばして前へ進む」、「原則 3：つまらなくなったら止める」という原則を多読 3 原則と定めている [4]。

3. 先行研究

足立らの先行研究では「SSS 式多読法に準じたテスト形式」、「項目応答理論を用いた能力推定」、「オンラインテスト」という特徴を有する Web アプリケーションを提案・開発を行った。

SSS 式多読法に準じたテスト形式とは試験内容が物語を読みながら設問に答える形式である。実施する試験では「物語を読むパート」と「設問に答えるパート」に分離された問題が出題され、受験者の能力推定が完了するまで出題が繰り返される。「物語を読むパート」では物語文に対して SSS 式多読法に則った読書を行わせるものであり、「設問に答えるパート」では受験者に設問を解答させるものである。

項目応答理論を用いた能力推定とは、試験の解答結果から確率論に基づいたテスト理論である項目応答理論を用いて能力推定を行う形式を指す。ERFPT ではさらに受験者の能力推定結果を用いて出題する問題を選出する適応型試験を実施することで能力推定精度の向上を目指している。

オンラインテストとはインターネットに繋がる環境さえあれば場所や時間に囚われず受験することが可能となる形式の試験である。そのため受験者の受験に対する負担を抑えられ、ERFPT ではテストの実施や採点・各受験者のレベルの遷移の管理等は全て Web アプリケーション側が行うことから教員の負担も抑えられる。

これらにより、受験者に対して少ない負担で適切なレベルを提示することが可能となった。しかし、先行研究での ERFPT ではシステムの全利用者を同一のサーバで管理しているため、複数組織で運用を行う場合において利用者の情報が他の組織に漏洩する懸念がある。また、教員が受講

生の学力の把握のために ERFPT の試験結果を確認できる成績管理機能や、多読用の能力レベルを提示する機能が実装されていなかった。そのため、教員は試験結果を参考に適切な指導を受講者に対して行うことができず、受験者についても推定された FRFPT における能力レベルを自身で多読用の能力レベルに変換を行う必要がある。実際に先行研究で開発された ERFPT に対する調査では ERFPT で提示されたレベルを参考に多読本を借りた受験者は少数であることがわかった。

4. ERFPT システムの再設計

4.1 Web アプリケーションの構成

3 章で挙げられた問題点のうち受験者の成績管理機能を実装するためには、図 1 に示すようにシステム利用者の情報と試験運用に必要な情報を別々のサーバで管理する必要がある。本報告では開発した ERFPT の構成はユーザ管理サーバとリソース管理サーバとなっている。ユーザ管理サーバでは各組織のシステム利用者のユーザ名、学籍番号、試験の成績といった情報が保存されており、システムの管理者や教授によって登録されている学生の情報が管理される。リソース管理サーバでは ERFPT の試験で出題する問題に関する情報やシステムを利用する組織の情報が保存される。この構成にすることでシステム利用者の組織を跨ぐ操作が行われなため、情報の流出の懸念を抑えることができる。

また、ユーザ管理サーバとリソース管理サーバとの通信は ERFPT のアクセストークンを用いた認証を行なっているため、リソース管理サーバへの不正アクセスに対しても対策済みである。

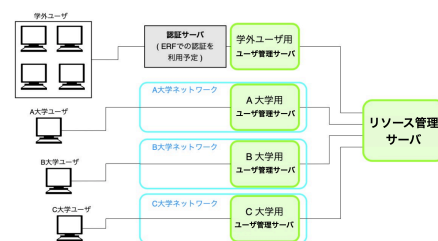


図 1 ERFPT のシステム全体の構成

各大学に「ユーザー管理サーバ」を設置し、それぞれのサーバが「リソース管理サーバ」にアクセス出来るような構成となっている。

4.2 仮想化技術を用いたシステム構築の簡易化

先行研究ではホストの環境に直接システムの構築を行っていたため、新たにシステムを構築する際に web サーバやデータベースなどのバージョンを揃える必要がある。そのためシステム構築にある程度の知識が必要であり、システム構築の際の労力や時間的なコストも大きい。

それに対して本報告では、開発したシステムの基盤に仮

想化技術の一つである Docker[5] というコンテナ技術を用いた。さらにこの Docker コンテナを Docker Compose[6] という仕組みを用いることでコンテナを一括管理している。このアプリケーションのコンテナ化とコンテナの一括管理によってアプリケーションの環境構築時に先行研究のようなコストを低減し、誰でも容易に ERFPT の環境構築が可能となった。

5. 能力推定結果の活用

5.1 LTI 認証と外部システムとの連携

本学も含め多くの教育機関では教育をサポートするツールとして LMS(Learning Management System)[7] を用いている。LMS の機能に LTI(Learning Tools Interoperability)[8] という外部ツールとの連携機能がある。本研究で開発した ERFPT では、LTI を利用し ERFPT 内部で認証処理を行わずにシステムを利用することが可能となった。さらに同じ LTI を利用しているシステムとの連携が可能であるため、先行研究では実現できなかった外部の学習管理システムとの連携も可能となった(図 2)。また開発したシステムでは外部の学習管理システムである本研究室で開発された ERS(Extensive Reading System)[9] と連携を行う設計となっており、ERFPT の試験結果を ERS 側に提供することや ERS での受験者の読書履歴を ERFPT で利用することが可能となった。この連携には HTTP のリクエストを用いているため、専用の API を実装するだけで ERFPT と ERS 以外の外部システムとの連携が可能となる。

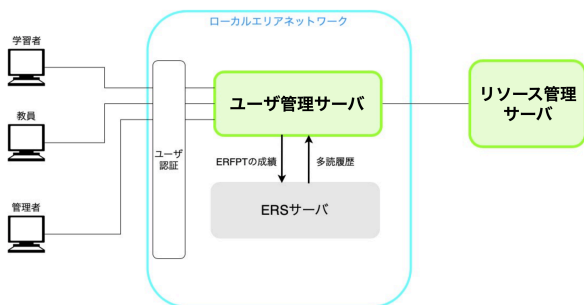


図 2 ERS との連携機能の概要図

各大学に設置されている ERS との連携が LTI での認証情報を用いて可能となった。主に ERFPT から ERS へは試験結果を送信し、ERS から ERFPT へは受験者の読書履歴データを送信する。またユーザ管理サーバは各組織の認証が必要なローカルエリアネットワーク内に構築されるため、ERFPT で独自にユーザ認証を行う必要はない。

5.2 尺度変換

本報告にて開発した ERFPT で推定された受験者の多読レベルは項目応答理論をもとに算出された値であるため、そのままでは英語多読学習に用いることは困難である。そ

こで ERFPT での推定値を変換して受験者に提示することで英語多読学習の支援に活用される仕組みを提案する。

英語多読学習では多読本の難易度を示す指標が幾つも存在するが、推定値の変換には ERS に保存される受験者の読書履歴を用いるため、本報告での変換後の値は YL[10] 相当の値となる。この値を YL 指標値とし、受験者の読書履歴に記録された読書本のレベルのトリム平均から求める。SSS 多読学習法は特性から学習者が自身の能力に適した多読本を大量に読むことを目的としているため、英語多読学習が進んだ学習者の読書履歴には学習者に適したレベルの多読本が多くなる。そのため英語多読学習が進んだ学習者ほど YL 指標値は学習者の多読能力に近い値となると考えられる。

本報告での提案手法は説明変数を「ERFPT の試験での能力推定値」、目的変数を「受験者の読書履歴から求められる YL 指標値」とする線形回帰分析により ERFPT での推定値を変換するモデルを生成するものである。

6. 評価実験

この章では 5 章で提案した ERFPT の試験での能力推定値を YL 指標値に変換する線形回帰モデルの精度について評価を行うために実施した評価実験の手順とその結果、さらに考察について記述する。

6.1 実験方法

実験は先行研究で開発された ERFPT で実施された試験結果から難易度等のパラメータが解析済みである 8 つの出題問題(大問)からそれぞれ設問数が 10 問のテストを生成し、それらのテストを受験者に解答させるものである。受験者の負担を考慮した結果、一度の試験で出題するテスト数を 4 つとした。2 種類の難易度が等しい試験を生成し、受験者 1 人あたり 2 回受験することとし、2 種類の試験の実施順は不正防止の観点からランダムとした。また、2 回の試験の受験間隔は 1 週間以内と短期間であり各試験で同じのテストは出題されないため 2 回の試験の間の試験準備の効果は考慮しないものとする。

試験はオンラインで受験する CBT(Computer Based Testing) であり、受験する場所と時間は自由かつ試験時間は約 30 分である。実験にて受験者の試験結果の収集が行われた後に、出題した項目に対する解答結果と読書履歴を用いて ERFPT の試験での能力推定値と YL 指標値を算出し、後述する尺度の変換についての評価を行う。

6.2 解析方法

本報告では評価実験の十分な実施期間が不十分であったため受験者の 2 回受験のうち 1 回目の受験結果から ERFPT の試験での能力推定値を求めた。また、線形回帰モデルの予測精度を評価する際に 10-fold クロスバリデーションを

実施する。

クロスバリデーションとは予測モデルの精度の評価の際にモデルの過学習を抑えるために用いる手法であり、モデルの学習に用いるデータを訓練用と評価用に分割し訓練用データでモデル生成を行い評価用データでモデルの精度を評価するものである。本実験では訓練用と評価用のデータの選出を 10 通り変更して線形回帰モデルの評価値の算出を行い、最後にそれぞれの選出における評価値の平均値を用いて線形回帰モデルの評価を行う。

モデルの評価の指標には RMSE(Root Mean Squared Error) と MAE(Mean Absolute Error), RMSE/MAE を用いた。RMSE と MAE は予測モデルの平均化された誤差の値を示すものであり、これらの値が小さいほど予測精度が高いことを表す。さらに RMSE は MAE と比べて外れ値の影響を受けやすいという特徴がある。RMSE/MAE は予測モデルが学習で用いたデータの大まかな特徴を表現できているかの指標であり、1.253 の値に近づくほど良い。

6.3 結果と考察

本実験では本校の英語多読を取り入れている講義の受講者に対して行われた。受験者のうち能力レベルの推定値と ERS の読書履歴から YL 指標値が取得可能であった受験者 263 名の能力推定値と YL 指標値を用いて評価を行った。実験で得られた ERFPT の試験での能力推定値と YL 指標値の分布を図 3, 図 4 に示す。また、実験で得られた本研究における説明変数と目的変数の対応を示す散布図を図 5, 線形回帰モデルの評価で 10-fold クロスバリデーションを行なった結果を表 1, 解析で生成した線形回帰モデルのうち RMSE が最小なものを図 6 に示す。

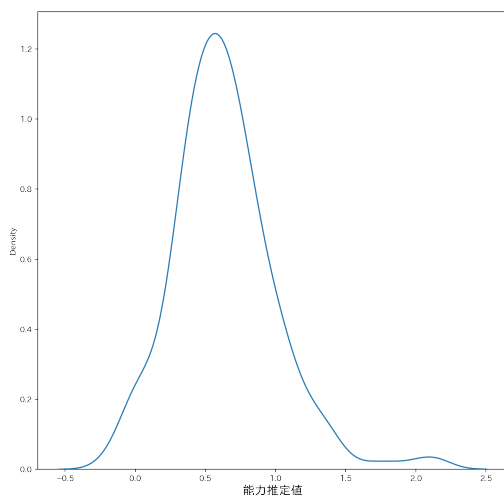


図 3 評価実験で得られた ERFPT の試験での能力推定値の分布 (カーネル密度推定結果)

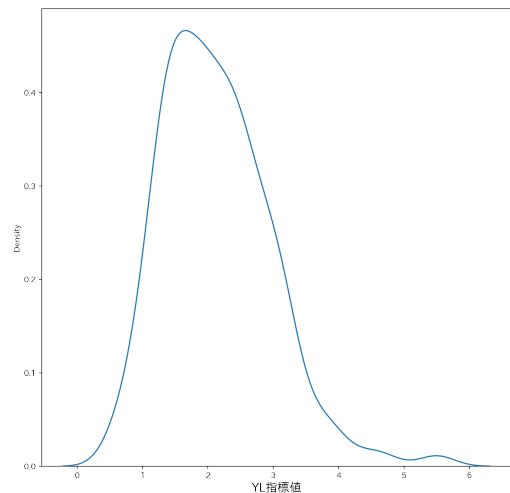


図 4 評価実験で得られた受験者の YL 指標値の分布 (カーネル密度推定結果)

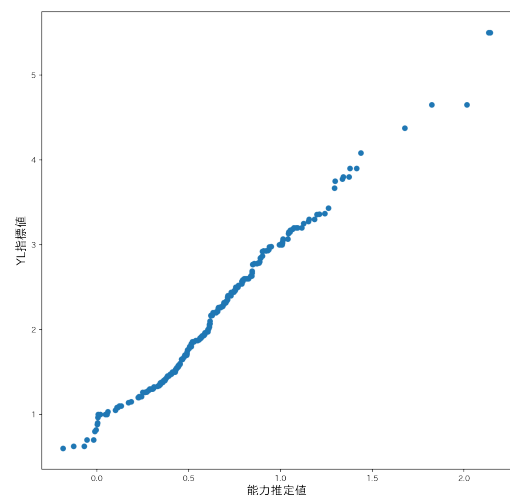


図 5 ERFPT の試験での能力推定値と YL 指標値の対応関係

横軸が ERFPT の試験での能力推定値、縦軸は YL 指標値となっている。

表 1 の各数値は 10-fold クロスバリデーションで得られた 10 通りの評価指標の値であり、ERFPT の試験での能力推定値から YL 指標値へ変換するモデルの RMSE の平均値は 0.1046, MAE の平均値は 0.08706, RMSE/MAE の平均値は 1.201 となった。

この結果から RMSE が MAE より大きく、RMSE/MAE が 1.213 よりも小さい。この要因として ERFPT による能力推定値で 0.4 以下の受験者が全体的に難易度の高い多読本を読んだ結果が影響している可能性がある。この原因としては表 2 から分かる通り、本学が蔵書する多読図書の

表 1 線形回帰分析の結果

試行回数	RMSE	MAE	RMSE/MAE
1	0.08759	0.07547	1.160
2	0.1096	0.09554	1.147
3	0.09635	0.08590	1.122
4	0.09201	0.07122	1.292
5	0.1497	0.09589	1.561
6	0.1019	0.08757	1.163
7	0.1184	0.09255	1.279
8	0.1083	0.1027	1.054
9	0.09013	0.07893	1.142
10	0.09241	0.08481	1.090

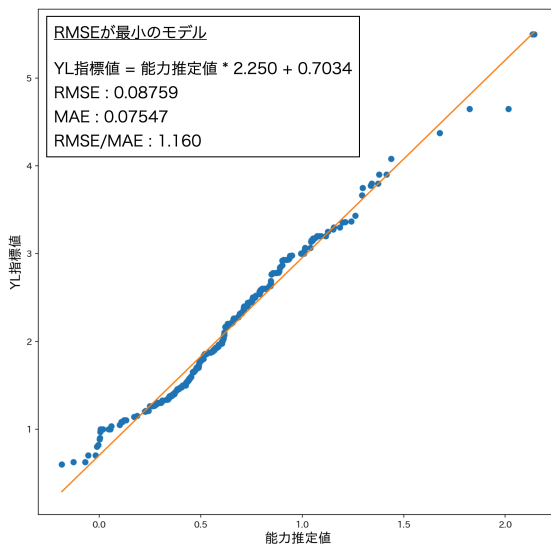


図 6 ERFPT の試験での能力推定値と YL 指標値の対応関係と生成した RMSE が最小の線形回帰モデル

横軸が ERFPT の試験での能力推定値、縦軸は YL 指標値となっている。

うち YL が 1 以下のものは数が比較的少ない。そのため学習者は興味を惹く多読図書が見つからず YL が 1 以上の多読本を借りた結果、多読学習が満足に行えていないことが考えられる。

さらに別の要因として ERFPT の試験での能力推定値では 1.4 以上の受験者の YL 指標値も線形回帰モデルと比べ誤差が大きいことが考えられる。この原因としては図 3, 図 4 から ERFPT の試験での能力推定値が 1.4 以上の受験者数が少ないことに加え、この水準の学習者は多読本を難易度よりも内容を重視して選出した可能性がある。

この評価実験の結果、YL は 0.1 刻みで設定される指標であり英語多読本に設定される YL の値の多くは 0.2 から 0.4 程度の誤差が設けられている。本報告で生成された線形回帰モデルの RMSE が 0.1 程度となったことから、英語多読学習で多読本を探す際の指標という用途において十分使用可能なものであると結論づけた。さらに能力推定値では 0.4 以下または 1.4 以上、YL 指標値では 1.3 以下または

3.5 以上のレベル帯の受験者について変換時の誤差が大きい問題については受験者のレベルの推定時にベイズ推定結果を信用区間とともに提示することで対処することが可能である。

表 2 本学が蔵書する英語多読本の YL 別冊数

YL	蔵書数
0 ~ 1	99
1 ~ 2	403
2 ~ 3	1026
3 ~ 4	1045
4 ~ 5	524
5 ~ 6	239
6 ~	111

蔵書数は受験者の大半が利用する図書館のものである。

7. まとめ

英語学習において多読学習が有効であり、多読学習を講義で取り入れることも多い。しかし多読学習を行う際に重要となる学習者の多読のレベルを把握することが困難であることから先行研究では学習者の多読能力を測る試験を実施するシステムである ERFPT を提案・実装した。

本報告では先行研究でのシステムの問題点の解決を目的として前半では他大学への導入を前提としたシステム構成について記述した。このシステムはユーザ管理サーバとリソース管理サーバで構成されており、それぞれのサーバは Docker によるコンテナ仮想化技術を基盤としていることが特徴である。さらに試験結果の活用のための外部システムとの連携機能について記述した。

後半では試験の実施によって得られた被験者の能力の推定値を多読学習で活用可能な値へと変換する手法の提案と変換処理についての評価実験を行い、結果について考察した。結果としては被験者母数から YL 指標値への変換の際には 0.1 程度の誤差が生じることがわかった。この誤差は英語多読学習において英語多読本を探す際の目安としては許容可能であるため被験者母数の推定の誤差が少ないとしたとき、ERFPT の試験の結果で受験者に提示される YL 指標値は英語多読学習において有効なものであると結論付けた。

問題点として ERFPT で推定された値の尺度の変換先が先行研究では ERF が定めた指標である ERFGrade[2] であり、本報告で開発したシステムも ERFGrade を受験者に提示することが求められることが挙げられる。この問題については今後の研究で ERFPT での試験による推定値から ERFGrade を算出する手法の検討を行う。

また、本報告で開発したシステムでは大学等の組織内のネットワーク上にユーザ管理サーバを設置し、組織のローカルエリアネットワークで実施される認証情報を利用して ERFPT のユーザの認証を行なっている。そのため、組織

に属さない一般ユーザがERFPTを利用できない。この問題を解決するために今後は一般ユーザ専用のユーザ管理サーバの設置とERFPT独自のユーザ認証機能の実装を行う必要がある。

謝辞 本研究を進めるにあたり熱心なご指導を賜った、信州大学大学院総合理工学研究科 工学専攻 新村 正明 准教授に厚く御礼申し上げる。最後に、本研究を行うにあたり、お力添えをしてくださった信州大学工学部及び大学院総合理工学研究科 新村研究室の方々に感謝する。

参考文献

- [1] 足立紘亮: 項目反応理論を用いた多読レベル測定のための適応型テストの開発と評価, 修士論文, 信州大学大学院工学系研究科 (2012).
- [2] Foundation, T. E. R.: The Extensive Reading Foundation Graded Reader Level Scale, http://www.robwaring.org/er/scale/ERF_levels.htm (2021/7/10 アクセス).
- [3] Group, S. E. R. S.: Sss 英語学習法のご案内, <http://www.seg.co.jp/sss/learning/index.html> (2021/12/25 アクセス).
- [4] 邦秀酒井: 快読 100 万語!ペーパーバックへの道, 筑摩書房 (2002).
- [5] Docker, I.: Empowering App Development for Developers — Docker, <https://www.docker.com> (2021/12/25 アクセス).
- [6] Docker, I.: Overview of Docker Compose, <https://docs.docker.com/compose/> (2021/12/25 アクセス).
- [7] Davis, B., Carmean, C. and Wagner, E.: The Evolution of the LMS : From Management to Learning (2009).
- [8] Severance, C., Hanss, T. and Hardin, J.: Ims learning tools interoperability: Enabling a mash-up approach to teaching and learning tools, *Technology, Instruction, Cognition and Learning*, Vol. 7, No. 3-4, pp. 245–262 (2010).
- [9] 武浩増田: SPA と認証認可プラットフォームによる多読学習支援システムの開発, 修士論文, 信州大学大学院 (2020).
- [10] Group, S. E. R. S.: YL and Word Count Tables, <https://www.seg.co.jp/sss/YL/index.html> (2021/5/30 アクセス).