

研究終了報告書

「読解に困難を抱える生徒を支援するための言語処理に基づくテキスト表示技術」

研究期間：2017年10月～2021年3月

(新型コロナウイルス感染症の影響を受け2021年9月まで延長)

研究者：松崎 拓也

1. 研究のねらい

研究代表者を含む研究グループでは2015年から、主として学校教科書から採った1~2文からなる短いテキストを用いた読解能力テスト(Reading skill test, RST)を開発し、小・中・高校生および成人のべ2万人に対し実施している(科研費 基盤A「テーラーメイド教育開発を支援するための学習者の読解認知特性診断テストの開発」)。その結果、課題および被験者の年齢によっては、事前の想定をはるかに超える割合で、教科書レベルの文が理解できない児童生徒が存在することが明らかになってきた。

本研究では、自然言語処理技術を応用し、個々の文の構造と内容、さらにユーザ毎の読解スキルに合わせ、電子デバイス上でのテキストの表示形態を自動的に設定することで、テキストの読解を支援することを狙う。また、RSTの実施の一部として、種々のテキスト表示形態が読解の正確さに与える影響を、数千人規模で調査する。また、「ふりがな付与」と「書字-音声の対応」あるいは「文節・句の間のスペーシング調整」と「統語構造の把握」といったように、テキスト表示要素とそれが支援する基礎的な読解スキルには自然な対応関係がある。これらを踏まえて評価実験の結果を解析することで、読解スキルの伸長モデルを構築することを狙う。

書体や太字などフォントの選択、行間隔の変更、ふりがなの付与、一部の文節間への空白の挿入、適切な位置での改行や、並列要素を揃えた字下げといったテキスト表示における様々な選択は、テキストの理解しやすさに極めて大きく影響することが予想させる。実際に、ディスレクシア支援の手段として、ハイライト・読み上げ同期ソフトの活用とともに、書体の改良を含めたテキスト表示上の選択の効果が研究されている。しかし、ディスレクシア以外の理由で読解に困難を感じる生徒にとっても、上記のようなテキスト表示の工夫による文章の易読化は、自力でテキストを読み解くための支援として有効だろう。読んで理解することの「中間ゴール」をテキスト表示形態によって可視化することで、暗黙知として捉えられがちな「読む」というスキルに具体的なイメージを与え、生徒が自らの力でテキストを読む力を育むことを狙う。

2. 研究成果

(1) 概要

数文～1段落程度のテキスト読解を支援するためには、構文構造に沿ったテキストレイアウトを行う事が最も有効だと予想した。2018年度および2019年度に延べ4500人の小中学生を対象としてRST読解課題におけるテキストレイアウトの効果を測定した結果、並列構造を含むテキストの一部で明らかな正の効果が観察されたものの、テキストレイアウトが有効であろうと事前に予想した問題においても、その大多数では明確な効果が観察されなかった。

また、テキストレイアウトとは異なる性質を持つ付加情報として問題文への振り仮名付与の

効果についての実験調査を行った。振り仮名の付与により、聞けば分かるが読めない語が認識できるようになる効果や、音声を通じたテキスト処理を近似することによる効果を期待した。しかし、小学校5年生から中学校2年生を対象とした実験では、どの学年でも振り仮名付与によるテスト成績の有意な向上は認められなかった(論文1)。

以上の結果を踏まえ、研究開始時の狙いであった、レイアウトや振り仮名といった付加情報による読解支援から、研究スコープを広げ、語彙知識と読解能力の関係の調査、一般学力と読解能力の関係の調査、問題文の言い換え・改変が読解成績に与える影響など、より多角的な観点から「なぜ読めないのか」を探る方向へと研究の中心を移した。

2019年度に行った調査では、読解テストと並行して語彙知識のテストを行い、語彙知識をテストした40語中35語で、それを知っている/知らない被験者の間で、当該の語を含む読解テスト課題に対する正答率が有意に異なることを明らかにした。また、同年の調査と一般学力テスト結果のクロス分析により、読解テスト成績と国語・数学の成績には中程度の相関があり、特に、中学校数学の問題のうち正答率が50%を切る問題において読解能力との関係が顕著であることを示した。

さらに、読解テストを用いた実験結果の分析に基づき、回答スピードおよびテスト成績の学年による伸長を定量化した。この結果を教科書テキストの定量分析結果と比較することで、小学校から中学校にかけての同一科目(理科)教科書の内容的な複雑さの増加にくらべ、読解能力の伸長ははるかに遅いことを明らかにした(論文3)。

新型コロナウイルス感染症の影響を受け6ヶ月間研究期間を延長し、教科書テキストに対する言語解析の高精度化を目的とする係り受け解析技術の改良に関する研究、および、板書形式の講義動画を非常に小さなデータサイズで作成するシステムの公開版の開発を実施した。前者においては、ニューラル言語モデルと係り受け解析器を結合することで得られる高精度化の要因について分析し、同格や並列関係が関わる特定の文型において大幅な解析精度の向上が見られることを明らかにした。後者においては、90分の講義動画で約4MBと、1/30程度のデータサイズでフルHDの動画と遜色ない品質を実現できることを確認し、編集機能を含めた公開用ソフトウェアの開発を行った。

(2) 詳細

研究テーマA「テキスト表示形態による読解支援」

2018年度に小中学生計450名、2019年度に小中学生計4000名を対象として、構文に沿ったテキストレイアウトが読解テスト成績に与える影響の実験調査を行った。特に並列構造を含む問題の一部で顕著な効果が見られた(図1)。しかし、全般にはテスト成績に対して正負両側の影響が見られ、いずれにおいてもその影響の幅は大きくなかった(図2)。

また、2017年度に小学校5年生から中学校2年生までの計200名を対象として、読解テスト問題に振り仮名を付与した場合のテスト成績に対する影響を調査した。振り仮名あり群となし群を比較したところ、回答時間に影響が見られたのは中学校2年生のみで、かつ、どの学年でも振り仮名あり群となし群の間にテスト成績に有意な差は見られなかった。この成果は認知科学分野のトップ国際会議であるCogSciで発表した。

研究テーマ B「読解能力と他の能力・知識との関係分析」

2019 年度に行った小中学生 4000 名を対象とした実験において、語彙知識と読解テスト成績の関係分析、テスト問題文の言い換えの効果の分析、および一般学力テスト成績との関係の分析を行った。特に読解テスト問題から抽出した語に対する語彙知識テストと、当該の語を含む読解テスト問題成績の関係の分析から、語レベルの知識の有無がテスト問題に対する正答率において 20%～30%程度の差として表れるケースが多数観察された。これは語レベルの知識の不足が読解能力の限界の一因となっていることを示すとともに、文中の語の意味を文脈から推測する力が弱いことを示していると考えられる。また、県レベルの共通学力テスト成績と読解テスト成績のクロス分析の結果から、読解能力と国語・数学の成績に中程度の相関(相関係数約 0.7)が見られ、特に中学校数学の正答率が 50%を切る問題において読解能力が高い群と低い群の間で正答率の大きな差が見られることが分かった。

洗濯用洗剤は

界面活性剤の種類と配合割合により、
石鹼、複合石鹼、合成洗剤に分けられる。

合成洗剤は

綿・麻・化学繊維用に 弱アルカリ性洗剤があり、
毛・絹用に 中性洗剤がある。

この文脈において、以下の文中の空欄にあてはまる最も適当なものを選択肢のうちから 1つ選びなさい。

中性洗剤は()である。

石鹼

複合石鹼

合成洗剤

弱アルカリ性洗剤

図 1:レイアウト付与によりテスト成績が向上した問題の例
(中学生正答率: レイアウトなし 50% → レイアウト有り 70%)

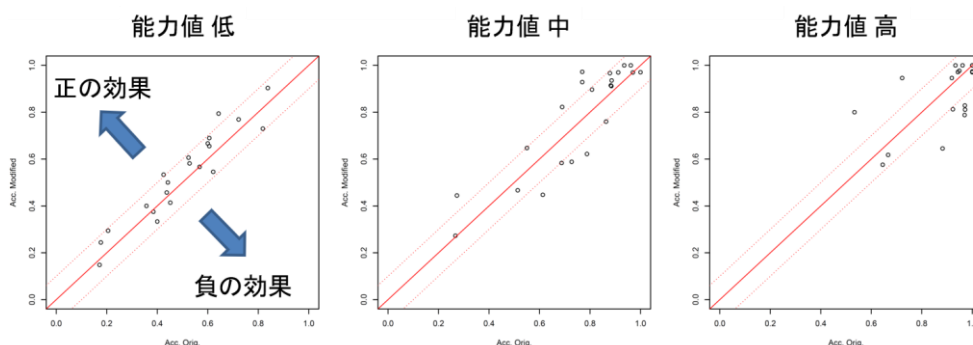


図 2:読解能力値 低・中・高群におけるレイアウトの効果
(横軸:レイアウト無し問題に対する正答率, 縦軸:レイアウト有り問題に対する正答率)

研究テーマ C「教科書テキストに対する言語処理技術の高精度化」

読解能力に関する分析のための基礎ツールとして、教科書テキストに対する高精度な言語解析技術が必要となる。教科書テキストは量が少ないため、新聞データ等の既存の訓練データに基づく解析器の分野適応が必須となる。少量の教科書データによる追加訓練の効果を分

析するために、解析器の内部状態をクラスタリングすることによる追加訓練前後の誤り例の分析を行った(論文2)。分析結果から、追加訓練の主要な効果は、分野特有の語句の特性の学習よりも、類似の文型に対する正しい構造の分布が分野間で異なることの学習によるものであるという示唆を得た。

コロナ延長期間中において、ニューラル言語モデル(BERT)を利用することで得られる係り受け解析処理の高精度化の要因について分析を行った。BERT を用いることで日本語係り受け解析の解析精度が向上するという事実がすでに知られている。本研究では、係り受けの品詞対や距離、また係り受け関係が関与する文型など様々な観点から解析精度向上の要因を分析し、同格や並列などを含む特定の文型の解析において大幅に精度が向上していることを明らかにした。

研究テーマ D「読解スキル伸長モデルの構築」

教科書テキストの分析と読解テスト結果の分析を総合することで、小学校に比べ中学校の理科教科書では概念定義の数が急増するのに対し、定義の理解を問う読解テスト問題に対する回答スピードと正答率は、小学校から中学校にかけて遥かにゆっくりとしか伸長しないことを明らかにした(図 3)。これは粗い概算ではあるものの、定義の理解に係る読解能力の限界によって中学校進学後に学習内容の積み残しが発生する可能性を示唆している。この成果は国際会議 CogSci で発表した。

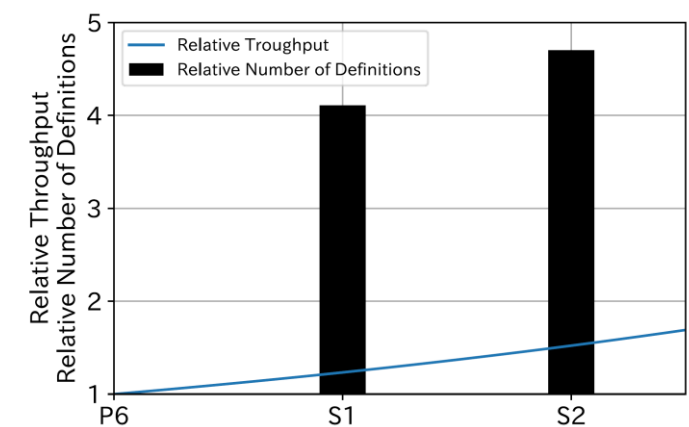


図3:教科書中の定義の数(棒グラフ)の増加と読解能力の伸長(青線)

3. 今後の展開

「さきがけ」研究期間の成果から、読解能力は語彙知識・文法知識・背景知識といった多レベルの要因が複合した複雑な現象であることが明らかになった。また、一般的な学力との相関や教科書を始めとする生活の中に存在するテキストの影響など、考慮すべき事象は多い。

これら多方面からの研究分析の蓄積が今後も必要である一方で、研究期間内に明らかになったように、語彙知識と読解能力の関係は明確であり、また語彙知識はテストおよび計量すること

が比較的容易な現象であることから、着実な基礎研究の方向として語彙知識と読解能力の関係の分析はより詳細かつ広範囲に展開して行くべきである。さらに、「さきがけ」期間内及び今後の研究成果を基に、読解能力の伸長のための有効な教育方策を見出すことが強く望まれる。

4. 自己評価

研究開始当初の目論見であった言語処理技術を用いたテキスト表示による読解支援は、実験調査の結果から有効性が確認できなかった。しかし、語彙知識の有無や問題文の改変による読解テスト成績の変化に関する調査分析、共通学力テスト成績と読解テスト成績の相関分析、さらに、教科書のテキスト分析と読解テストに関する実験調査結果を総合した定量的な知見を基に現在の教科書テキストの問題点を指摘した点など、当初目的のための実験調査体制をベースに社会的にもインパクトのある成果を生むことができたと考える。

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数: 3件

1. Teiko Arai, Kyosuke Bunji, Naoya Todo, Noriko H. Arai, Takuya Matsuzaki. **Evaluating Reading Support Systems through Reading Skill Test.** Proceedings of the 40th Annual Cognitive Science Society Meeting. 2018 年. pp. 100 – 105.

小学校 5 年生から中学校 2 年生までの約 200 名の被験者を対象に、読解能力テスト課題に振り仮名を付与した場合とそうでない場合の結果の比較を行った。テスト課題として Reading Skill Test の問題を用いた。課題に対する反応時間(回答時間)およびテスト成績に対する統計分析の結果、中学 2 年生の被験者のみ、振り仮名を付与した場合の反応時間に有意な増大が見られた。また、どの学年においても、振り仮名の有無でテスト成績には有意な差が認められなかった。

2. 原 拓也, 松崎 拓也, 横野 光, 佐藤 理史. **誤り例のクラスタリングによる係り受け解析器のドメイン適応効果の分析.** 自然言語処理. 2020 年. 27(1). pp.3-30.

新聞で訓練された言語解析器を教科書に適用する場合など、訓練時と運用対象のテキストドメインが異なると解析精度は一般に低下する。これを防ぐ代表的な手段である追加訓練の効果を分析した。具体的には、解析モデルの内部状態のクラスタリングによって追加訓練前後の誤りを自動的に分類し、定性的・定量的な分析を行った。結果から、追加訓練の主要な効果は、類似した文型に対する正しい構文構造の分布がドメイン間で異なることを学習することであるという示唆を得た。

3. Teiko Arai, Takuya Matsuzaki, Hidenao Iwane. **Appraising Science Textbooks through Quantitative Text Analysis and Psychometric Results of Students' Reading Skills.** Proceedings of the 42nd Annual Meeting of the Cognitive Science Society. 2020 年. pp. 3240 – 3246.

小学校および中学校の理科教科書における概念定義の数を調査し、中学校教科書では定義の数が急増することを明らかにした。さらに、定義の理解を問うタイプの読解能力テスト課題の結果を分析し、回答スピードおよび正答率から推定される、定義を理解する能力の平均

的な伸長が，教科書における定義の数の増加に追いついていないことを指摘した。

(2) 特許出願

研究期間累積件数:0 件(特許公開前のもも含む)

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

1. 原 拓也, 松崎 拓也, 佐藤 理史. **係り受け誤り埋め込み表現のクラスタリングによるドメイン適応の効果の分析**. 情報処理学会自然言語処理研究会, 2018 年.
2. 原 拓也, 松崎 拓也, 横野 光, 佐藤 理史. **係り受け誤り埋め込み表現のクラスタリングによる複数のドメイン間での追加訓練の効果の比較分析**. 人工知能学会全国大会(第 33 回). 2019 年.
3. 松崎 拓也, 東中 竜一郎. **AIは「人を真似る」のか**. UP. 2019 年 5 月号. 東京大学出版会.
4. 相原理子, 石川香, 藤田早苗, 新井紀子, 松崎拓也. **コーパス統計量と読解能力値に基づいた単語の既知率の予測**. 言語処理学会第 27 回年次大会. 2021 年.
5. 青木拓磨, 原田裕文, 三浦大輝, 新井紀子, 松崎拓也. **読解能力テストに対するニューラル言語モデルを用いた自動解答及びその結果の分析**. 言語処理学会第 27 回年次大会. 2021 年.