

児童が書いた文章の内容把握を支援する インタフェースの提案と有用性の検証

松原未和^{†1} 加藤直樹^{†2}

概要：教員は児童が書いた文章をもとに自身の授業を振り返り、授業構想を練ったり、グループ活動の編成を考えたりすることがあるが、記述内容を確認・把握し、これらを行うことは容易ではない。そこで、我々は「誰が」「何を」書いたのか、似た意見や考えをもっている児童は誰かを把握することを支援するインタフェースを提案してきた。本稿では、そのインタフェースの概要とともに、予備的評価に基づいて行なった改善と、教員がねらいとした内容を記述できている児童を見つけるときのツールの有用性を検証した評価実験について述べる。評価実験の結果から、特定キーワードを文章中で書いている児童を見つけたり、全体像を把握したりする場面で本ツールが活用できる可能性を示すことができた。またツールを効果的に使用できる場合、紙を使用して該当児童を見つけるよりも要する時間が短くなることも明らかにした。

Proposal and usefulness validation of interfaces to support understanding of texts written by children

MIWA MATSUBARA^{†1} NAOKI KATO^{†2}

1. はじめに

2020年度より小学校で施行されている学習指導要領の改定に合わせて、教職員向けに授業改善のための参考資料が公表され、その中で児童生徒の学習評価についての項目が記載された。資料の1つである「児童生徒の学習評価の在り方について(報告)」[1]には学習評価の在り方として、

- ・教師の指導改善につながるものにしていくこと
- ・児童生徒の学習改善につながるものにしていくこと
- ・これまで慣行として行われてきたことでも、必要性・妥当性の認められないものについては見直していくこと

と記載されている。また、文部科学省が提示する「新しい学習指導要領の考え方」[2]では、主体的・対話的で深い学びの実現について、『『主体的・対話的で深い学び』の視点に立った授業改善を行うことで、学校教育における質の高い学びを実現し、学習内容を深く理解し、資質・能力を身につけ、生涯にわたって能動的(アクティブ)に学び続けられるようにすること』とある。これらのことから、児童の学習状況に応じて授業改善を行うことの重要性はますます高まっていくと言える。

児童の学習状況の確認や学習の成果として活用できるものの一つに授業を振り返って短時間で書く短い文章(本稿では授業感想文と記す)が挙げられる。そこで我々は、児童が書いた授業感想文の内容を把握することの支援を目

標に、「誰が」「何を」書いたのか、似た意見や考えをもっている児童は誰かを把握することを支援するインタフェースを提案してきた[3]。本稿では、[3]で述べた概要に加え、小学生が書く文章の解析結果を適切に表現するための対話インタフェースについて述べるとともに、予備的な評価実験を元に加えた改良と、提案したインタフェースの有用性を検証するために行った評価実験について述べる。

2. 記述内容の把握を支援するインタフェース

2.1 基本コンセプト

我々が提案するインタフェース[3]では、高瀬らによる、教員は記述内容の確認を行う際に、使用されているキーワードの確認をした後にキーワードの使用法を確認し、最後に解答全文を確認するという手順を踏んでいる[4]との主張に沿って、

- (ア) 授業感想文で使用されている単語の確認
- (イ) 授業感想文中での単語の使用法の確認
- (ウ) 書かれた授業感想文そのものの確認

ができれば記述内容の確認手順が再現できると考え、授業感想文に含まれる「単語」に着目し、単語の使用頻度や品詞の可視化、用いられることが多い単語やその文章を特徴づける単語の確認、児童が書いたもとの文章の閲覧を可能にする。

また、個体と個体の接続関係の集合によって表現される

^{†1} 東京学芸大学大学院 教育学研究科
Graduate School of Education, Tokyo Gakugei University

^{†2} 東京学芸大学
Tokyo Gakugei University

情報を可視化するにはネットワーク状のマッピング（以下、ネットワーク図とする）が採用される [5]。加えて、近くに座る児童同士でグループを組み、意見をまとめる機会や授業の冒頭に異なる視点から書かれたいくつかの授業感想文を発表し、活動に活かす機会があることから、授業感想文の内容を座席配置図に反映させることで授業感想文の大別や内容の把握が容易にできるようになると考えられる。

そこで本インタフェースでは、授業感想文に含まれる授業で用いられた単語に着目した文章解析によって得られた結果をネットワーク図と座席配置図へマッピングする。加えて、これらのマッピングによる可視化において、教員が注目したい点を反映した結果を表示、確認できるようにするため、対話的（インタラクティブ）に操作できるようにする。

2.2 対話インタフェースの概略と追加設計

本節では、基本コンセプトを実現する対話インタフェースについて、[3]で述べた基本部分の概略を述べた後、新たに追加したノードの新規追加と編集インタフェースの設計を述べる。

2.2.1 ネットワーク図を用いた記述内容の可視化

ネットワーク図を用いた可視化では、授業感想文から得られる情報の把握を簡易化するために、授業感想文に含まれる単語とその単語を使用した児童、およびそれらを結びつける情報を可視化する（図 1）。また、注目したいノードを選択することで、選択した児童が使用した単語や選択した単語を使用した児童だけを抽出して表示できるようにする（図 2）。児童名に注目した場合は、どのような文脈で単語が使用されたのかを確認できるように、その児童が記述したもとの文章を表示する。

2.2.2 座席配置図を用いた記述内容の可視化

似ている意見や考えをもっていたり、同じような授業感想文を書いたりした児童を教員がまとめて確認できるようにするために、児童のグルーピングを行い、座席配置図へのマッピングで表現する（図 3）。

グルーピング方法としては、指定した観点をもとにしたグルーピング、文章の類似性をもとにしたグルーピング、類似度の高さをもとにしたグルーピング、共起関係にある語をもとにしたグルーピングの 4 種類を提供する。

2.2.3 ノードの新規追加と編集

前記ネットワーク図を用いた可視化を実装したところ、誤記のある文章やひらがなが多い文章では、文章を単語に分解する処理が正しく行われず、単語ノードに表示される単語名が意図しないものとなったり、正しい表記であれば表示されるはずのノードが表示されなかったりするケースが生じることがわかった。加えて同じ単語を使用しているも、表記に用いる文字種（ひらがなや漢字など）が児童によって異なると別の単語ノードとして表示されてしまうことや、本来 1 語で表示したい単語が 2 語以上に分解されて

しまうケースがあることがわかった。

そこで、これらを修正するため、ノードに表示される単語名の変更や同義の単語ノード同士の結合、単語ノードの新規追加を可能にする。

例えば、同じ対象（メダカ）について書いてある授業感想文でも児童によって「メダカ」と「めだか」のように表記が異なっていた場合、ネットワーク図上ではこの 2 つの単語は別のノードで表現されてしまう。このような場合に前述したノードの結合を行う。結合したいノードの選択は、ネットワーク図に表示されている全ての単語ノードから選択できるようにするが、結合対象となる可能性が高い語を見つけやすくするために、2.3 節で定義する同義語と準同義語を優先的に提示する。

2.3 解析手法の設計（児童が書いた文章に適した分散表現モデルの作成）

ここでは、2.2.3 項で述べた同義語を求める際の設計を述べる。なお、授業感想文をネットワーク図に表示するための解析、児童の意見をグルーピングするための解析については [3]を参照されたい。

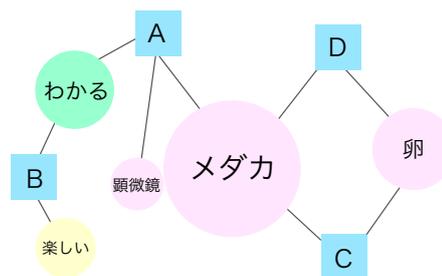


図 1 ネットワーク図による分析結果の可視化

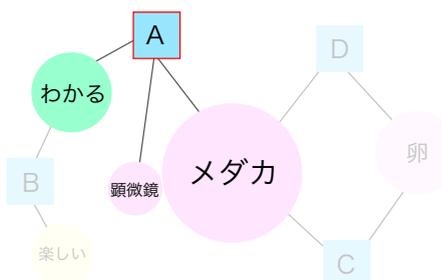


図 2 児童 A に注目したときの抽出表示



図 3 座席配置図による分析結果の可視化

前 2.2.3 項で述べたノードの結合を行う際、選択した単語ノードの結合候補となる別の単語ノードを見つけるために、ある単語ノードの同義語として、その単語と分散表現が近い語を定義する。また、準同義語を同じ漢字が含まれる語として定義する。

同義語を求める際に使用する分散表現モデルは様々な大学や企業、団体、個人で作成・公開されている [6, 7, 8, 9] が、これらは小学校の学習内容に対応させて作成されたものではないため、学習指導要領の範囲外の用語が同義語となる可能性があり本研究での使用には適していない。そこで、本研究では、国立国語研究所が公表している教科書コーパス語彙表 [10]をもとに分散表現モデルを作成する。

教科書コーパス語彙表は、小学校・中学校・高等学校の全学年・全教科(科目)の教科書1種ずつを対象に、その教科書に記載されている語彙とその読み、品詞、活用形、校種ごとの度数・使用率・特徴度、教科ごとの度数・使用率・特徴度、初出学年などの一覧になっている。そこで、教科書コーパス語彙表に記載されている語彙を見出し語とする Wikipedia の記事を検索し、そこに記載されている文章を学習対象テキストとして分散表現モデルを作成することとした。

まず、学習対象テキストを収集するために、教科書コーパス語彙表に記載されている語彙を見出し語とする Wikipedia の記事をスクレイピングした。教科書コーパス語彙表に記載されている地名および人名は全てカタカナで表記されていたことから、一般的に漢字で表記する地名・人名については CradleExpress1.1 [11]を使用し、漢字表記を調べ、書き換えた後にスクレイピングを行った。

ある単語について同義の単語を見つけ出したいことから、収集した学習対象テキストに対し、形態素解析を行い文章を分かち書きにした。その後、Wikipedia で使用されている注釈記号を削除し、Word2vec [12]を用いて単語からベクトル空間を作成した。学習対象テキスト内で似たような文脈で使用される単語の単語ベクトルは、生成したベクトル空間内で近い位置に配置されるため、同義語を見つけることが可能になる。

3. 可視化ツールの試作

提案したインタフェースの有用性を検証するために、それを実装したツールを試作した。

本ツールは、ハードウェアキーボードを有するパーソナルコンピュータをターゲット環境として開発を行った。フロントエンドに JavaScript とそのフレームワークである Vue.js を用いた。また、ネットワーク図の描画には JavaScript ライブラリである D3.js を、座席配置図の描画には jQuery プラグインである jQuery-Seat-Chart.js を用いた。バックエンドには Python3.8 とフレームワークの Flask を使用した。教員が実施した授業に関するデータや授業感想文を保存す

るデータベースとして Google 社が提供する Firebase Cloud Firestore を用いた。授業感想文や授業の情報の登録、および解析結果の表示には Web ブラウザを使用した。

4. 評価実験

児童が書いた授業感想文に対して、教員がねらいとした内容を記述できているか否かに基づいて分類を行い、ねらいとした記述ができている児童を見つけるとき、提案したインタフェースを実装した本ツールが有用なのかを、時間、正確性、分類をする際に読んだ文章量を元に検討することを目的とした評価実験を行なった。

4.1 実験内容

被験者は、本学に通う小学校での教育実習経験がある学生 14 名とした。

小学生が書いた授業感想文の文意を変えず文章を整えたデータセットを 4 つ作成した(表 1)。データセット 1「メダカの誕生」は、小学校第 5 学年理科で取り扱う動物の誕生に関する内容であり、設問「メダカの誕生の単元で学んだことを書きましょう」に対し、該当単元で学んだことについて記述されている。データセット 2「スーホの白い馬」は、小学校第 2 学年国語で取り扱う物語文であり、設問「スーホの白い馬を読んで思ったことを書きましょう」に対し、主人公スーホやスーホが育てた白馬、登場人物のとのさまについて児童が感じたことが記述されている。データセット 3「お手紙」も、データセット 2 と同様に小学校第 2 学年国語で取り扱う物語文である。設問「どうして二人は書いてあることがわかっているお手紙を長い間待っていたのでしょうか」に対し、物語文に記載されていた事実をもとにした理由や自分の考えを含めた意見が記述されている。データセット 4「動物園の獣医」は、小学校第 2 学年国語で取り扱う説明文である。ペンギンが誤って飲み込んでしまったボールペンを獣医が取り出すことが書かれた場面を学習し終えた後の設問「感想を書きましょう」に対し、ペンギンの行動や獣医が行った治療についての感想が記述されている。

ツールの説明と操作に慣れてもらう時間を設けた後に、用意したデータセットに含まれる授業感想文を分類する評価実験を行った。被験者には、データセットを表 1 に示した分類基準に従って 2 つのグループに分類(二値分類)ないし、3 つのグループに分類(多値分類)してもらった。データセット 1, 2 が二値分類、データセット 3, 4 が多値分類となる。なお、多値分類の分類基準はその授業を行った教員にヒアリングをして作成した。

被験者をランダムに 2 つのグループに分け、片方のグループはデータセット 1, 3 については紙に書かれた文章を見ながら分類してもらい、データセット 2, 4 はツールを用いて分類してもらった。もう片方のグループはデータセット 1, 3 の分類でツールを使用し、データセット 2, 4 で

表 1 データセットの内容と分類基準

番号	単元	設問	分類基準
1	メダカの誕生 (小5, 理科)	メダカの誕生の単元で 学んだことを書きましょう	メダカについて書いているか
2	スーホの白い馬 (小2, 国語)	スーホの白い馬を読んで 思ったことを書きましょう	白い馬についての感想を書いているか
3	お手紙 (小2, 国語)	どうして二人は書いてあることが わかっているお手紙を 長い間待っていたのでしょうか	1) 事実を元にした理由になっている 2) 登場人物の気持ちを理由にしている 3) 1, 2 を踏まえて自分の気持ちを書いている
4	動物園の獣医 (小2, 国語)	(該当段落の) 感想を 書きましょう	1) ペンギンがボールペンを飲み込んでしまった ことのみを書いている 2) 薬を使って助けられること, 獣医さんが助けた ことを書いている 3) 上記以外

は紙を用いて分類を行った。このとき、分類にかかった時間を計測した。多値分類を行う際には、事前になぜその基準にしたのか、基準に該当するかを考えるとときにはどのような単語や表現に注目しているのかを被験者へ共有した。各データセットの分類終了後、判断に迷った授業感想文にはマークをつけてもらった。

評価実験終了後に事後アンケートに回答してもらった。ツールを使用して分類を行った場合と紙を使用して分類を行った場合それぞれについて、分類を行う際に全員分の文章を読んだか、文章を分類することは大変だったかを、「とてもそう思う・とても当てはまる」を2点、「全くそう思わない・全く当てはまらない」を-2点とした5段階で回答してもらった。また、ツールを使用した分類と紙を使用した分類を比較して、気づいたことや感想などを自由記述形式で回答してもらった。

4.2 結果

評価実験を行い、得られた結果を本節に記す。なお、データセット1の分類基準を間違えていた被験者が1名いたため、データセット1のみ被験者13名の結果となっている。

4.2.1 所要時間

各データセットの分類を終えるまでに要した時間について、ツールを使用して分類した場合と紙を使用して分類した場合ごとの平均値と95%信頼区間を図4に示す。

4.2.2 一致率

あらかじめ分類の模範回答となる正解データを用意しておき、正解データとの一致率を算出した。被験者が分類した結果と正解データとの一致率について、ツールを使用した場合と紙を使用した場合をそれぞれ縦軸と横軸にとった散布図を図5に示す。

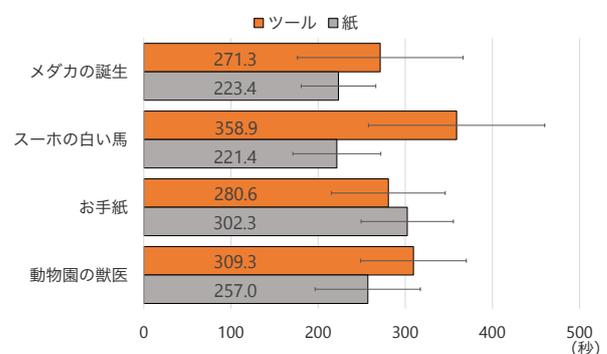


図 4 分類に要した時間の平均 (秒) と 95%信頼区間

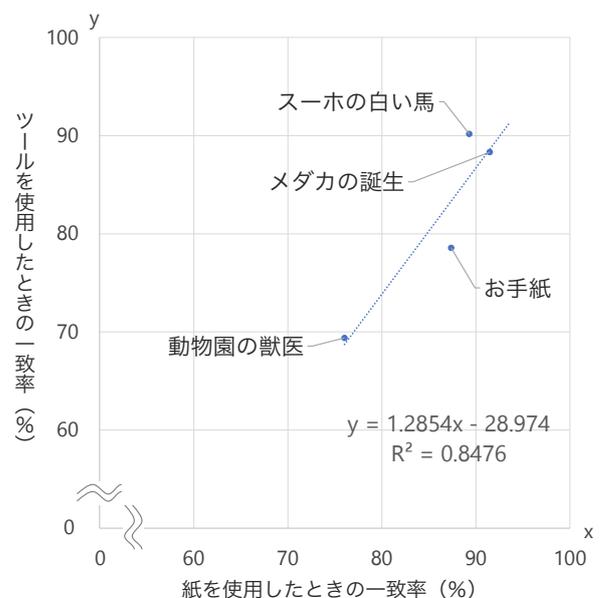


図 5 紙を使用した場合とツールを使用した場合の分類結果と模範回答との一致率の平均値

4.2.3 主観的評価

二値分類と多値分類で、ツールを使用して分類を行った場合と紙を使用して分類を行った場合のそれぞれに対して、分類を行う際に全員分の文章を読んだか、文章を分類することは大変だったかについて主観的評価の回答の内訳と平均点および標準偏差を集計したものを図 6、図 7 に示す。

4.2.4 自由記述回答

得られた自由記述回答を次にまとめる。

- ・ ツールに使い慣れていないため、紙でやる方が楽だと感じたが、記述内容が一目でわかるツールは使い慣れたら便利になるのではないかと思います。
- ・ 紙だと子供それぞれ異なるノートのページをめくったり、向きを揃えたりなどの手間がかかるが、ツールを使用すると簡単なクリック操作のみで済むので楽だと感じました。また、AND や OR 表示を用いることで、全ての文章を読む必要がなく時間が短縮できた。
- ・ ツールだと、最初にどのくらいの人が理想的な文章を書

けているのかがすぐ分かるなと思った。紙だとどうしても一枚ずつ読むことになる。また、紙だと全部読み終わった後全体像をうまく把握できなかったが、ツールだと全体を把握するのが容易だなと感じた。

- ・ キーワードが入っているかいないかについては、ツールを用いることでかなり作業が捗るなと思いました。
- ・ ツールの機能を最大限に活かせなかったのが、紙で行った分類の方がやりやすいなと感じました。
- ・ ツールの使い方がわかってくるとやりやすいと思った。
- ・ 紙での分類では分類を進める中で、一度分類した紙を選択できないように端に置いたりできるが、今回のツールではもう既に分類が完了している番号（人）についても表示されてしまったため、今回の分類の手法とする表に数字や記号を書いていく作業だと難しいという風を感じました。
- ・ 35 人程度なら紙が速いかもしれないが、100 人程度ならツールが速くなってくると感じました。

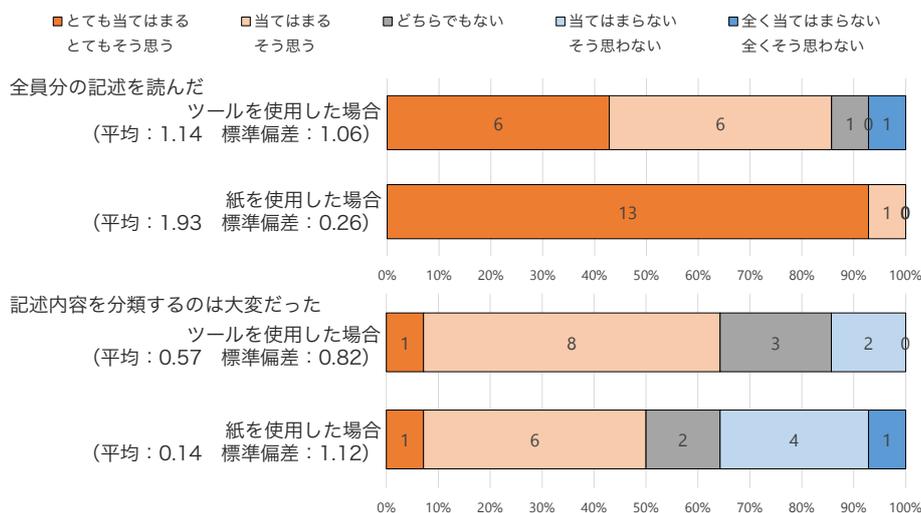


図 6 二値分類についての主観的評価

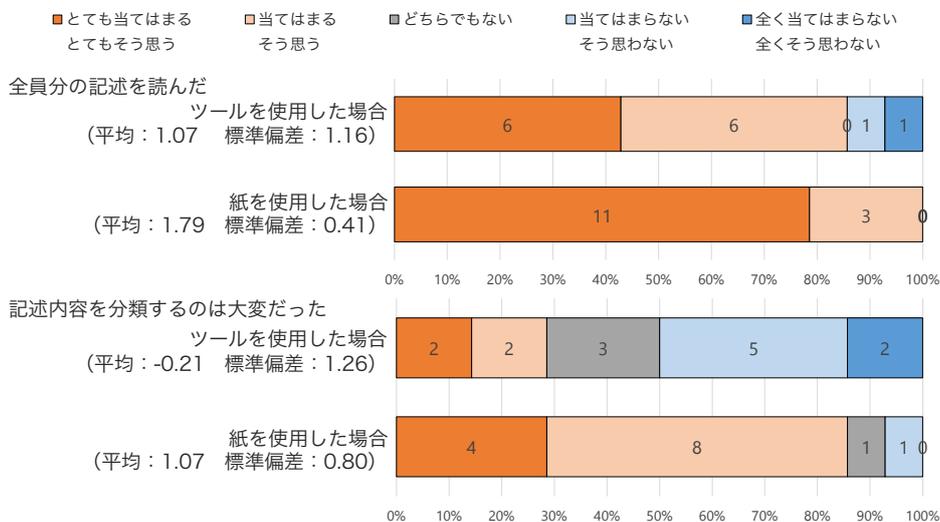


図 7 多値分類についての主観的評価

- ・二値分類の場合は文章量が長すぎなければ紙の方が早く分類できた。だが、多値分類では、判別に時間がかかるので、ある程度ツールで単語ごとにまとまっているとわかりやすかった。
- ・ツールを利用すると、文章の分類を語句やキーワードに着目しながらできるので、効率的に分類を行うことができるように感じた。

4.3 考察

4.3.1 所要時間とツールの使用方法についての考察

平均所要時間は、データセット3を除き、紙を使用して分類した方が速いという結果になった(図4)。

ここで、ネットワーク図や座席配置図の各機能を用途に合わせて使い分けたり組み合わせたりしていた被験者を「ツールを効果的に使用できていた被験者」(以下、効果的使用群、それ以外を非効果的使用群)とし、それぞれの群、および図4で示した被験者全員が紙を使用して分類した場合の平均所要時間を比較すると、データセット2を除き、効果的使用群がツールを使用して分類した場合が最も短いことが明らかになった(図8)。また自由記述回答には、「記述内容が一目でわかるツールは使い慣れたら便利になるのではないか」、「キーワードが入っているかないかについては、ツールを用いることでかなり作業が捗る」、「ツールの使い方がわかってくるとやりやすい」、「ツールを利用すると、文章の分類を語句やキーワードに着目しながらできるので、効率的に分類を行うことができるように感じた」といった回答が寄せられた。これらのことから、本ツールは操作に慣れることで、教員がねらいとした内容を書いている児童を効率的に見つけることに有用である可能性が示唆された。

データセット2の結果をネットワーク図で表示すると図9のようになる。他のデータセットのネットワーク図(例

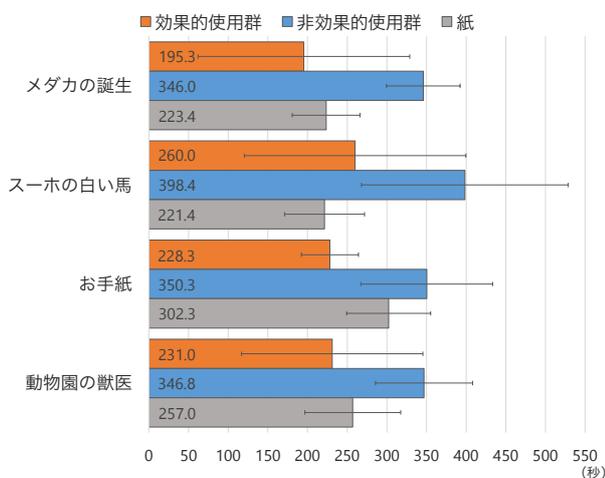


図8 群ごとのツールを使用した場合と全被験者が紙を用いた場合の分類に要した時間の平均(秒)と95%信頼区間

としてデータセット3の結果を図10に示す)と比較して、データセットに含まれる授業感想文の半数以上が使用していた単語ノードが登場人物に関するノードに偏っている一方、使用人数が3以下の単語ノードが多くあり、全体としてノード数が多いことがわかる。紙を使用した分類では、重ねて置いてあるものの上の方からや、たまたま目についたものから見ていくが、ツールを使用する際は、最初にどこに注目すれば良いか目移りしてしまうことで、分類に取り掛かるまでに時間がかかってしまう。これが分類の所要時間が長くなった原因の一つだと考えられる。また、今回の評価実験の事前説明では、「この操作を行うとネットワーク図や座席配置図の表示がこのように変わる」という操作をベースにした説明を行ったが、「〇〇をしたいときには△△という操作が良い」というように何が実現されるのかをベースにした説明を行うことで、各機能の使用場面が把握しやすくなり、分類の所要時間が短くなる可能性があると考えられる。

4.3.2 分類の種類とツールの有用性についての考察

全員分の記述内容を読んだかを尋ねた主観的評価では、紙を使用して分類した場合は、二値分類と多値分類ともに8割弱〜9割強の被験者がとても当てはまると答えた一方、ツールを使用して分類した場合には、とても当てはまると答えた割合は4割強に留まった(図6, 図7)。このことから、ツールを使用することで全員分の文章を全て読まなくても授業感想文の分類が行える可能性があることが示唆された。

記述内容を分類するのが大変だったかの主観的評価で、二値分類については、ツールを使用して分類した場合と紙を使用して分類した場合それぞれの回答の傾向に大差はなかった(図6)。一方、多値分類については、ツールを使用した際に記述内容を分類するのは大変ではなかったと回答した被験者が半数おり、紙を使用した場合より多くなっている(図7)。多値分類においては、本ツールを使用することで、紙を使用して分類を行うときと比較して簡単にこなせる可能性がうかがえる。自由記述回答に「二値分類の場合は文章量が長すぎなければ紙の方が早く分類できた。だが、多値分類では、判別に時間がかかるので、ある程度ツールで単語ごとにまとまっているとわかりやすかった」という意見が寄せられていることから、本ツールは教員が児童の記述内容をもとに、児童をいくつかのグループに分ける、というような作業を行う際に有用である可能性が示唆された。

4.3.3 一致率についての考察

平均一致率の散布図に示す回帰直線($r=0.92$, $R^2=0.8476$)から、紙で分類したときの平均一致率が高ければツールで分類したときの平均一致率も高い傾向にあることがわかった(図5)。また、二値分類の方が多値分類よりも一致率が高い傾向にあることも明らかになった。分類基準が、ある

に触れているので1と回答し、判断に迷ったことを示すマークをつけていた。

以上のことから、授業感想文の分類を行ったときの一致率の高低はツールの有無ではなく、書かれている文章の読みやすさや分類のしやすさに関係していると考えられる。

5. おわりに

本稿では、児童が書いた授業感想文の内容を把握することの支援を目標に、[3]で提案したインタフェースの概要を述べた後、予備的な評価実験を元に加えた改良と、提案したインタフェースの有用性を検証するために行った評価実験について述べた。

評価実験の結果から、本インタフェースが、教員がねらいとした内容を書いている児童を効率的に見つけること、児童の記述内容をもとに児童をいくつかのグループに分けるような作業を行うことに有用である可能性が示唆された。また、ツールを使用することで全員分の文章を全て読まなくても授業感想文の分類が行える可能性があることがわかった。

今回はデータセットを作成する際にもとにした授業感想文の学年・教科が偏っていた。そのため、複数学年・教科のデータセットで検証していくことが今後の課題として挙げられる。また、今回の評価実験は学生を対象に行なったため、同様の評価実験を現職の教員に対しても行い、システムの有用性を検討していきたい。

参考文献

- [1] 文部科学省. 児童生徒の学習評価の在り方について (報告).
- [2] 文部科学省. 新しい学習指導要領の考え方.
- [3] 松原未和, 加藤直樹. 授業感想文の内容把握を支援するインタフェースの提案と改善のための評価. コンピュータと教育研究会. 2021, Vol.2021-CE-160, No.1, pp.1-10.
- [4] 高瀬治彦, 川中普晴, 鶴岡信治, 森田直樹. 記述式小テストの解答群の分析手法—解答群からのキーワードの自動抽出—. コンピュータ&エデュケーション. 2013, Vol.34, pp.46-49.
- [5] 伊藤貴之. 意思決定を助けるための情報可視化技術. コロナ社, 2018.
- [6] “日本語 Wikipedia エンティティベクトル”.
http://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/, (参照 2022-1-5).
- [7] “word2vec の学習済み日本語モデルを公開します”.
<https://aial.shiroyagi.co.jp/2017/02/japanese-word2vec-model-builder/>, (参照 2022-1-5).
- [8] “Word vectors for 157 languages”.
<https://fasttext.cc/docs/en/crawl-vectors.html>, (参照 2022-1-5).
- [9] “Pre-trained word vectors of 30+ language”.
<https://github.com/Kyubyong/wordvectors>, (参照 2022-1-5).
- [10] 国立国語研究所特定領域研究「日本語コーパス」言語政策班. 教科書コーパス語彙表
- [11] “CradleExpress 1.1”. <https://cradle.ninjal.ac.jp/>, (参照 2022-1-5).
- [12] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient Estimation of Word Representation in Vector Space. 2013, In ICLR Workshop Papers.