

学生レポートの自動添削システムによる学習支援 (2)

Study Assistance using Automatic Proofreading System (2)

共同研究メンバー

○中村その子*、出原至道* (○代表、執筆者)

キーワード：ライティングスキル、校正、自発的学習

Keywords：writing skill, proofreading, self-directed learning

1. 背景・目的・手法

本研究は、2019 年度から継続する共同研究である。2019 年度の成果については、既報である [1]。まず、あらためて簡単に背景、目的、手法を説明する。

学生がエントリーシートや志望動機書など、ある程度フォーマルな形式・内容の文書を作成する際に指摘してほしいと感じている項目には以下のようなものが挙げられる。

文字・単語レベル

- 句読点の間違いや誤字脱字
- ことばの誤用 (らぬき言葉、助詞、慣用句など)
- 「です・ます」調と「である」調の混在
- 送り仮名・表記・漢字の間違い
- 接続詞の選択
- 差別語、禁忌語などを使っていないか
- 話し言葉の使用
- ケアレスミステイクの指摘

文レベル

- 文 (脈) 中での慣用句の使用
- よじれ表現
- 冗長表現
- 適切な段落づくり (改行)

文脈・文章レベル

- 文章の構成や言いたいことの順番
- 内容の重複などがなく論理構成がしっかりしているか
- 言いたいことがしっかり表現できているか

* 多摩大学経営情報学部 School of Management and Information Sciences, Tama University

上述のケアレスミスや文字・単語レベルの間違いについては、該当部分が「間違っている」という指摘を受けて、学生自身が自力でそれを修正できれば、指導の効率がよくなる。教員は文レベル、もしくは文章レベルの指導（加えて、学生が何を間違ったのかわからず自力で修正することができない文字・単語レベルの間違いの指導）に集中することができ、文章作成指導の効率が上がると考えられるからである。

また、教員として気になり、また指摘しても学生がどこが間違っているのか気づきにくい間違いとして「考える、思う、感じるなどが用いられている複文のよじれ」がある。

例：「わたしが感じているの（こと）は作者は日本の農村の歴史を文化的側面からとらえていると思います」

本研究では、これらの間違いについて、即時・自動的に指摘するシステムを開発することによって、学生の自発的な学習効果を高めることを目的とする。文章校正システムとして RedPen [2] を採用した。RedPen は、多摩大学の卒業生である伊藤敬彦氏によって 2014 年に開発が開始された文書フォーマットチェックツールである。日本語・英語・韓国語などの様々な言語への対応と、設定ファイルによる柔軟なルール設定などの特徴を持つ。

2. 2019 年度までの成果と課題

2019 年度には、RedPen を学内サーバで動作させ、Google Document から RedPen と通信して文章表現をチェックしコメントを表示するシステムを開発した。これによって、学生が簡単に校正結果を確認できるようになった。実際に本システムを学生に利用させ、その使用感をアンケートによって評価することは、今年度の課題として残された。

3. 2020 年度の成果

3.1 システムからのコメントと学生の反応

RedPen のコメントに対する学生の理解度・自発的修正の可能性を調べるため、ゼミ生 32 名を対象として調査を行った。調査では、サンプルの文章とそれに対する RedPen のコメントとを提示し、「どう直せばよいかわかる」「いわれていることは分かるが、どう直せばよいかはわからない」「何が問題なのかわからない（何を言われているかわからない）」「これは問題ないと思う（指摘そのものが間違っていると思う）」の 4 つの選択肢から適当なものを選ばせた。

RedPen のコメントには、重大なレベルの順に Error, Warn, Info があり、レベルごとに出力を抑制することができる。このうち、Error, Warn について、学生の回答を示す。

例文 1：

外国のコマーシャルは、近年は、商品よりも雰囲気や重きを置いていて、そのため、商品の詳細の実情の意味はわからなく、何となくどんなものかしかが分からないが、アメリカは、ユーモアを大事にする文化であり、企業は商品をコマーシャルで説明して、理解してもらおうという考えより面白い humormessage から買おうと考えてもら

おうとしている傾向にある。

RedPen からの指摘：

- [Warn] 文長 ("173") が最大値 "120" を超えています。(図 1)
- [Info] 助詞「の」が連続しています: " ~商品の詳細の実情の~ "
- [Info] 一文に二回以上利用されている単語 " コマーシャル " がみつかりました
- [Info] 一文に二回以上利用されている助詞 " て " がみつかりました。
- [Info] カンマの数 (9) が最大の "3" を超えています。
- [Info] アルファベット単語の前にスペースが存在しません
- [Info] アルファベット単語の後にスペースが存在しません。

例文 2：

また、外国のコマーシャルは日本人からすると過激に感じる人が多いが、外国では、自社の製品が優れていることを示すために、(もともとアメリカは競争社会なので、他社の製品を強く非難したり攻撃したりすることや、自社の製品の優れたところをどんどんアピールすることに抵抗感が少ないのだろう) コマーシャルで他社を挑発するような内容を取り入れることが多いです。

RedPen からの指摘：

- [Error] パラグラフの開始部分が不正です "ま" (図 2)
- [Warn] 括弧が長すぎます。(図 3)
- [Error] 不正な日本語のスタイルを発見しました " です " (図 4)

例文 3：

そのため、日本の企業は「他社と比べて・・・」ではなく、「消費者にとってどれだけ良いか」をコマーシャルで PR する必要があると思っています。

RedPen からの指摘：

- [Info] 不正な表現 " 思っています " がみつかりました。
- [Error] 不正な日本語のスタイルを発見しました " ます " (図 5)

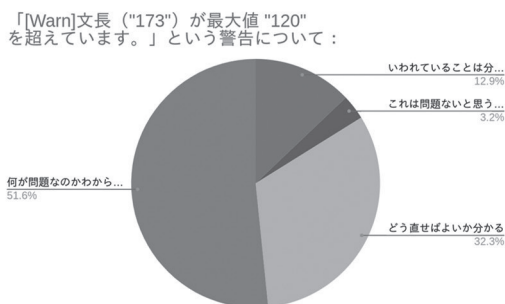


図 1：文長

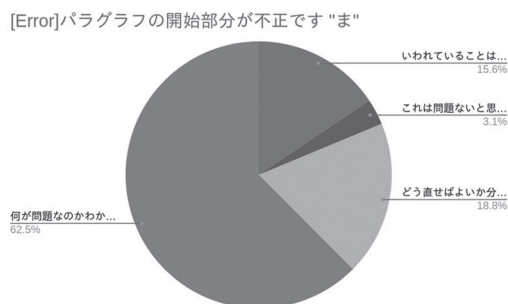


図 2：段落の開始

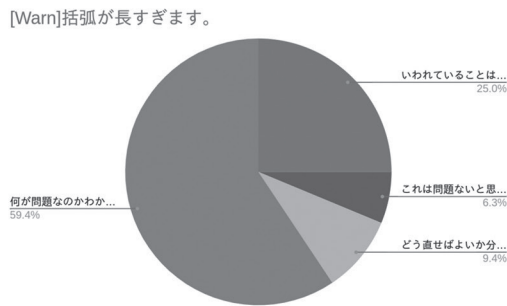


図 3 : カッコ

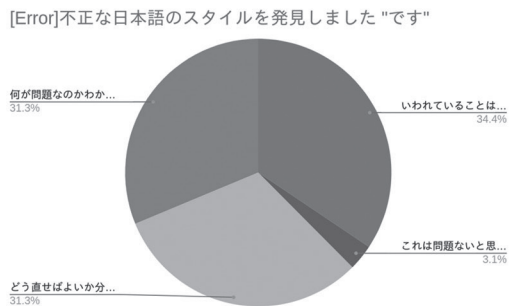


図 4 : 「ですます」

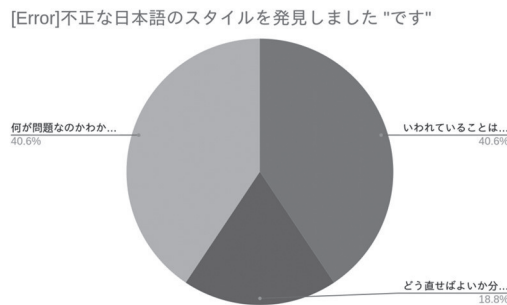


図 5 : 「ですます」

図 1 の「文章が長い」警告については、半数以上の学生が「何が問題なのかわからない」としている。論文・レポートなどで、文章をダラダラと長く書くことに対して抵抗感がない、あるいはむしろ、長い文章が良い文章であるという誤解があると推察される。

図 2 では、エラーメッセージが「パラグラフの開始部分が不正」と記述されているため、6割の学生が、指摘を理解できていない。これについては「段落の最初は 1 文字空けなければならない」などの表現とすることで改善できる。

図 3 の「カッコでくくられた部分が長すぎる」警告については、6割の学生が指摘を理解できていない。カッコによる挿入文への抵抗がないと推察される。

図 4・図 5 の「ですます」については、半数以上が指摘の内容を理解している。2つの指摘に対して学生の回答が異なることについては、例文 2 が「である」と「ですます」が混在しているのに対し、例文 3 では「ですます」で統一されているためであると考えられる。今回のシステムでは、2019 年度の本研究で導入した「である」を正調と固定するオプションを指定しているため、文章全体が「ですます」で統一されていたとしても Error と判定する。論文・レポートでは「である」が正調であることをあらかじめ指導しておくことで、理解が容易になる。

3.2 システムの改良

システムを運用している過程で、「見れる」がエラーとして検出されていないことが判明した。これは、RedPen がチェックに使っている辞書 MeCab に、可能動詞として「見れる」が登録されているためである。「見れる」「見れた」などの表現が、「一段動詞+れる」ではなく「一般動詞」と判定されるため、誤りとして発見されなかった。「見れる」以外にも、このような可能動詞が登録されている可能性があるため、プログラム側で「辞書に登録されていても誤り」

であると判定する必要がある。このため、データ構造とプログラムに以下の改良を加え、改善提案 [3] を行った。なお、「ら入れ」の誤りもこの判定の過程で検出が可能となったが、現段階では未実装である。

終止形が「れる」で終わり、かつ「られる」で終わらない動詞について、可能動詞としての正誤を以下の手順で判定する。

1. 元の動詞の終止形の「れる」を「る」に置換した文字列を作成する（「見れる」→「見る」、「切れる」→「切る」）。これによって、可能動詞の元になった動詞を得る。
2. この動詞の活用が「五段活用」でなければ、元の動詞が辞書に載っていたとしても、誤りであると判定する。（「見る」(上一段活用) → 「見れる」は誤り 「切る」(五段活用) → 「切れる」は正しい)

4. 考察

本研究によって、学生が理解しにくいメッセージや、あらかじめ指導しておくべきポイントが示唆された。今後、学生が RedPen からのメッセージを理解できるものとしたうえで、RedPen を採用したときの文章指導の効率化について評価したい。

また、システムとして、主述のねじれや、主語の流れの検出について、実装可能性を検討していきたい。

[1] 出原至道, 中村その子, 「学生レポートの自動添削システムによる学習支援」, 経営・情報研究 多摩大学研究紀要 (2021), 25, pp.219-222

[2] <https://github.com/redpen-cc/redpen>

[3] redpen-cc/redpen への Pull Request #881 <https://github.com/redpen-cc/redpen/pull/881>

