

大学生の日本語文章力向上を目的とした
提出文章の要改善箇所指摘システム

三浦 智美 中島 潤
北海道情報大学

The improvement point indication system for advancement of
Japanese writing ability for college student.

Tomomi MIURA and Jun NAKAJIMA
Hokkaido Information University

平成28年 3 月

北海道情報大学紀要 第27巻 第2号別刷

〈論文〉

大学生の日本語文章力向上を目的とした
提出文章の要改善箇所指摘システム

三浦智美* 中島 潤†

The improvement point indication system for advancement of
Japanese writing ability for college student.

Tomomi Miura* Jun Nakajima†

要旨

大学生の日本語能力が問題視されるようになって久しいが、授業等の課題として学生が提出したレポートについて、日本語文章として問題がある箇所を細部にわたりチェックし、コメントを残して返却することは現実的に難しい。そこで本論文では、大学生の日本語文章作成能力を向上させるために、文章の改善箇所を機械処理により自動的に指摘することで、自発的に文章の推敲を促すためのシステムを提案し、またその有効性を評価するために提案したシステムを試作し検証したので報告する。

Abstract

A decline of Japanese writing ability of Japanese college student is considered to be a problem. In this article, we propose a system for making the elaboration of a report in voluntarily basis to improve Japanese writing ability of college students, and verified its effectiveness by experimental systems.

キーワード

日本語文章 文章作成能力 言語処理 教育システム

* 北海道情報大学大学院経営情報学研究科, Graduate School of Business and Information Systems (Grad. of BIS), HIU

† 北海道情報大学経営情報学部准教授, Associate Professor, Dept. of BIS, HIU

1. はじめに

1-1 本研究の背景

大学生の日本語力の低下が問題視されるようになり、大学生の「書く力」をめぐる、1990年代後半から、「文章表現」「日本語表現法」などといった日本人学生に対する日本語科目を設置する大学が増え始めた^[1]。2011年7月に読売新聞より発表された「第4回『大学の實力』調査(623校)^[2]」によれば、47%の大学で、「日本語を学び直す特別授業」が行われ、レポートの書き方などの文章作法の授業が添削指導という形で実施されている^[3]。

しかし、添削指導を行う授業は、学習者の文章作成能力の向上には有効であるが、評価を行う教員の負担は大きく、特に学生数が多いと、個別の学生に対して指導すること自体が困難になる^[4]。実際、本学の「日本語表現」の授業でもレポートを書く場が前後期合わせて3、4回あり、教員は1度に100人以上から提出されるレポートを添削しなければならない、1人のレポートにつき20分以上の時間をかけて添削を行っている場合もある。また、添削も1度で終わらず、再提出といった形となることも多いため、教員にかかる負担がうかがえる。

そこで本研究は、本学学生の日本語文章力向上のために、文章の改善箇所を機械処理により自動的に指摘することで、自発的に文章の推敲を促すためのシステムを提案し、またその有効性を評価するために提案システムを試作し検証した。

1-2 本研究の目的

本研究では、本学学生の日本語文章力向上を目的とし、そのために2つの方法を選んだ。1つは、学生が自ら改善箇所を確認し自発的に推敲をすることで、同様の問題

を繰り返させないようにすることである。もう1つは、教員が従来、基本的な形式・文法の誤りといった箇所の指摘に割かれていた時間を軽減することである。これにより、標題に沿った論理展開や内容をよりわかりやすくするための文章構成の指導・採点、学生の意識問題に対する改善の取り組みなど、学生への個別対応が取りやすくなると考えた。

本研究における日本語文章力とは、レポートを執筆するにあたって、語彙・語法などといった基本的な日本語表現を理解し、論理的でわかりやすい文章を書くことを指す。

本稿では、第2章で大学生の日本語力の現状とその現状解決に向けた先行研究について述べる。第3章では、本学学生の日本語文章力向上に向けたシステムの提案を行う。第4章では、提案システムを試作するにあたって、本学学生の提出レポートにおける文章の質低下要因を分析・明確化し、質低下の主要な原因となる指摘項目の指摘指標を統計的手法によって分析を行い決定した。さらに、第5章では、Webアプリケーションによるシステムの試作開発について述べ、第6章では、本学教員によるレポートの添削結果と試作システムの指摘を比較することにより、指摘の精度による提案の有効性を評価した。

2. 提出文章の要改善箇所を指摘するシステムの提案

本提案システムは、インストールや動作環境を整える煩わしさを避けるためにWebアプリケーションとした。学生がレポートを完成させ提出する前にWebを介して、レポートの形式・文章のルールといった誤りの指摘をする。またそれと同時に、一文が

長すぎるといった文章の質の低下要因となる項目も指標を基に指摘を行う。これにより、学生は指摘を基に自ら文章を推敲し、また教師の負担を軽減することで、学生の日本語文章力の向上を図る。ただし本提案システムは、Tomarigi¹⁾のような細部までのチェックをするようなものでなく、本学教員へのヒアリング調査によって、優先的に挙げられた項目を含む基本的な文章の形式・文法の誤りと、一般的に読みづらいとされる文章の主要原因を中心に指摘するものとした。またここでいう指標とは、本学学生の基準に合わせたものとするために、本学学生の提出レポートを基に文章の質を低下させている要因を分析・明確化し、決定したものである。

本提案システムを試作するにあたって、まず文献調査 [5][6][7][8][9] 及び教員へのヒアリング調査を通じて、文法・形式的な誤りといった基本的な指摘項目を洗い出した。またそれと同時に、本学学生の提出レポートについて、文章の質低下要因を分析し明確化し、質低下の主要な原因となる指摘項目の指摘指標を統計的手法によって分析を行い決定した。これらを基に、システムの試作を行い、その有効性について評価・検証を行った。

3. 学生のレポートの分析

3-1 分析対象

本研究では、提案システムを試作するにあたって、文章の質低下の要因を分析・明確化し、指摘する指標を本学学生に合わせ

たものに決定するために、計 377 部の本学学生のレポートの収集を行った。

3-2 分析結果

本研究では、文法・形式的な誤りといった要因以外の、文章の質を低下させている要因について分析を行った。本研究ではその要因を、本学教員の評価、添削・指摘コメントを基に、指摘率の高いものを主要要因として 5 項目に分けた。

3-2-1 内容の把握しづらさ

本項目は、「一文の長さ」「主語と述語の不一致」といった文章の意図を伝えるにあたって最も核となる部分を取り扱う。

- 一文の長さ

一文の長さは、読み手に文意を伝えるにあたって重要な役割を担っている [10]。長い一文は、それだけで正しい文章で書いてあっても意図を読み取ることの困難さ、後述にもある主語と述語の不一致、文意が読み取れなくなるほどの文章崩壊を引き起こす。ただし、内容の把握しづらさと一文の長短とは本質的に関係がないという意見もある [11]。事実、一文が短ければ必ずしも良いというわけでもなく、極端に短い文章は、箇条書きが並んでいるような単調な文章になりがちである。しかし、この問題においては、書き手の実力に依存するため、書き手が日本語に通じていない場合、短い方が無難であるとされている。

本分析においての一文の長さの指摘は、全体の 3 割以上とする中、高評価を得ているレポートは一切の指摘が見受けられなかった。また、評価が低くなるほどされる人数も多くなっており、E ランクでは 7 人のうち 5 人が指摘を受けている(表 1)。更に、指導教員からのコメントにおいても、正確な文意を取れるものとするために、できる

¹⁾ Tomarigi (校正・推敲ツール) は、入力文章に対して、校正・推敲対象となる指摘箇所をマークし、指摘理由や修正候補を提示する事で、文章校正の支援を行う。

だけ短い文を書くように指摘されているものもあった。

表 1. 一文の長さの指摘人数と指摘率

| 評価 | A | B | C | D | E | 計 |
|------------|---|---|-------|-------|-------|-------|
| 人数 | 3 | 3 | 22 | 8 | 7 | 43 |
| 一文の長さの指摘人数 | 0 | 0 | 7 | 3 | 5 | 15 |
| 指摘率 | 0 | 0 | 0.318 | 0.375 | 0.714 | 0.349 |

● 主語と述語の不一致

主語と述語の不一致は、一般的にねじれ文と呼ばれるものであり、主語と述語の関係が意味的に繋がらなくなるものである。また、ねじれ文の中には、本分析においても、一文の長さの指摘と同等数の指摘が見受けられ、評価の低い文章は、殆どが指摘を受けていた。また、指摘の中には、「強いて述べるなら～と述べる。」「そういったことを考えていくと～と考える。」といった文頭と文末で同じ表現を使っている二重表現が見受けられた。更に、文頭と文末の不一致に関しては、理由を述べるにあたって使われる「なぜなら」が、「～からである。」と文末で締めくくられていない者がいた。

3-2-2 読みづらさ

本項目では、「改行の過不足」「漢字の使用率」といった、文章を読むにあたって文法が正しくても読みにくいと判断される部分を取り扱う。

● 改行の過不足

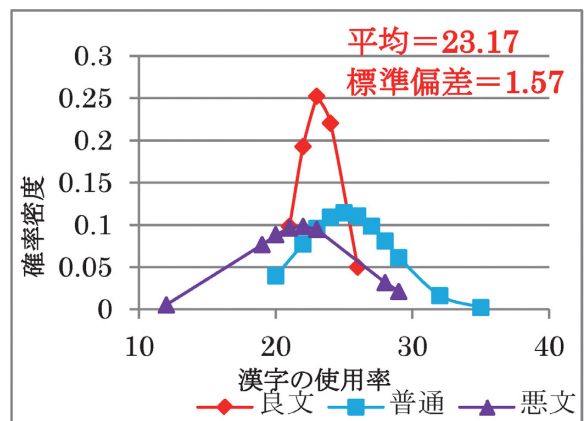
段落は、文章全体を構成する要素である [12]。段落のない文章は、内容の切れ目を判断しづらく、文章の展開や主要な情報、考えが把握しづらくなる。また、その逆で、メールと同感覚で句点がかかるたびに改行が行われる段落の多すぎる文章も、内容の切れ目が判断しづらい。

ただし、本分析に使用したレポートにおいては、手書きの文章であるためか極端に改行された文章も段落わけが全くされていない文章も殆ど見当たらなかった。また、1字下げにおいても、手書きの文章では段落ごとに正確に行われていた。

● 漢字使用率

文章を読むにあたって、極端にひらがなだけが羅列している文章は、たとえ文法がすべて正確であったとしても、多くの場合非常に時間を要する。また、漢字使用率が少ない文章は、稚拙文にもみられやすい。本研究では、漢字使用率についても文章の質を落とす要因として考えた。

そこで評価ごとにおける漢字使用率の分布を出した(グラフ 1)。グラフから、良文は普通文に比べ全体的に漢字の使用率が低いことになっているが、これはサンプル数の差によるものであると推測した。しかし、普通文と悪文を比べてみると、漢字の使用率の違いが明らかにわかるものとなっている。



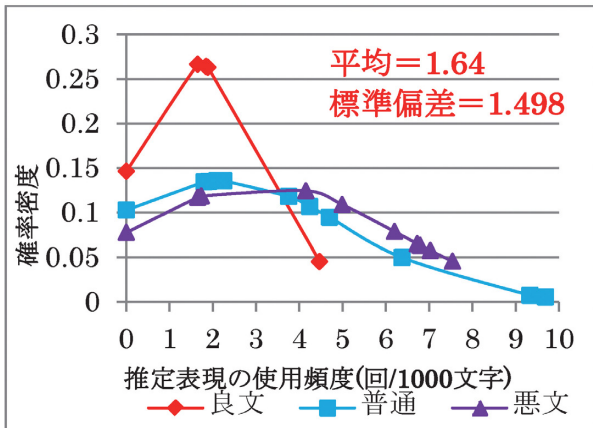
グラフ 1. 評価ごとにおける漢字の使用率

3-2-3 根拠のなさ

次に、文章の質を低下させる要因として挙げられるのが、「思う」「だろう」などの推定表現を使用することによる根拠のなさである。また、一般的にも、レポートでの

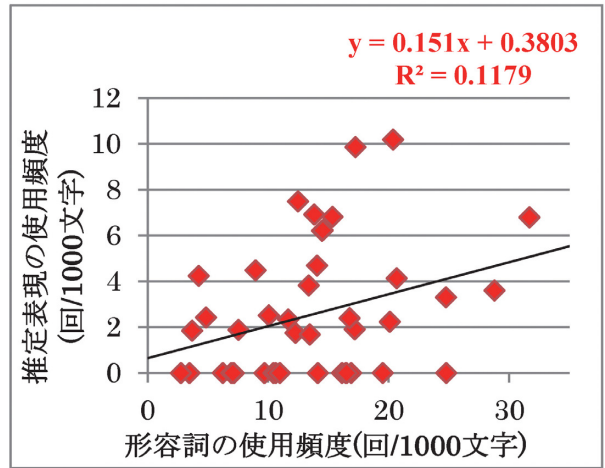
「思う」「感じる」といった表現は、主張を弱めてしまうため好ましくないとされている。

そこで、提出レポートの評価ごとにおける推定表現の使用率について分析した。ここでは、文章量による差異も考え、1000文字ごとにおける推定表現の使用頻度を出した(グラフ2)。その結果、やはり良質な文章であるほど、推定表現の使用頻度は少なかった。普通文と悪文における差異が殆どみられないのは、本レポートにおける課題が自分の体験談を踏まえて述べたものであるためと推測した。



グラフ2. 評価ごとにおける推定表現の使用頻度

また、形容詞は、情緒的表現が多いことから、推定表現の使用頻度と形容詞の使用頻度における相関関係について調べた(グラフ3)。その結果、有意水準5%において、推定表現の使用頻度と形容詞の使用頻度における相関関係はほぼないことがわかった。

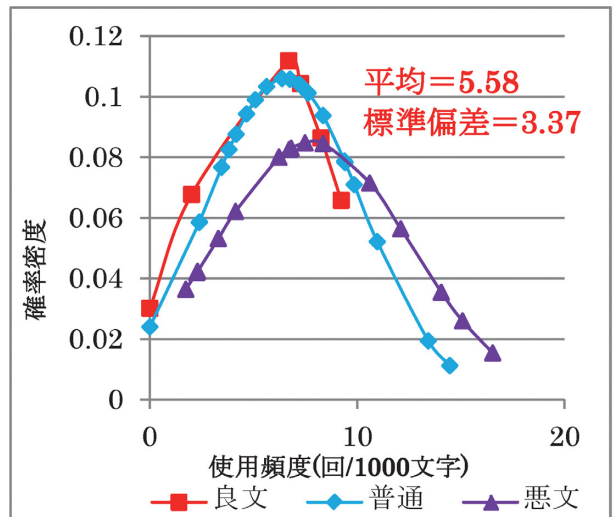


グラフ3. 推定表現と形容詞の使用頻度の相関

3-2-4 曖昧さ

文章を書くにあたって、同じ言葉を使わないというものがある [13]。それが最も顕著に現れているのが英語であるが、日本語においても文章作法としてそのように捉えられている。この同じ言葉が並ばないようにするための方法の一つとして、指示代名詞の使用が挙げられるが、指示代名詞の乱用は文意を曖昧にしてしまう。

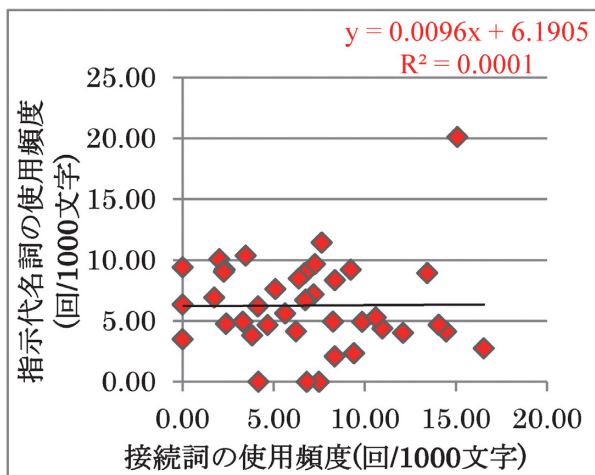
そこで評価ごとにおける指示代名詞の使用頻度について分析した(グラフ4)。その結果、良文であるほど指示代名詞の使用頻度は少ないことがわかった。ただし、良文と普通文における差異があまり見られないことから、サンプル数が増えた場合、双方間の差異はない可能性が考えられる。



グラフ4. 評価ごとにおける指示代名詞の使用頻度

しかしながら、指示代名詞は、接続詞の代用としても使用することができるという意見もある [14][15]。つまり、指示代名詞の使用頻度が少なくなることで、逆に接続詞の使用頻度が増えている可能性がある。接続詞が極端に少ない文章も理論構成が崩れるため問題であるが、多用は稚拙な文章に見えてしまう。

そこで、指示代名詞の使用頻度と接続詞の使用頻度について、相関関係があるか分析した(グラフ 5)。その結果、有意水準 5%において、有意差がなかったため、指示代名詞の使用頻度による接続詞の使用頻度の差があるのかどうか判定できなかった。



グラフ 5. 指示代名詞と接続詞の使用頻度の相関

また、曖昧表現の要素として、「色々な機能がある」「様々な面がある」といった具体例がない表現や「～〇〇等」「～〇〇とかがある」といった具体的な例が 1 つしか挙げられていないのに、複数要素あるような曖昧表現が見受けられた。

3-2-5 文脈の繋がらなさ

最後に、本項目では、本学学生が文章を書くにあたって、不用意に用いやすい語句を洗い出した。不用意に用いやすい語句は、「たしかに」「ある意味」「いわゆる」と文頭に意味のない不要な装飾がされる場合と、

「具体的には」と書いてあるのに具体性がない、例文に対する例文といった「例を挙げると」の多用、「次に」といった繋がらないのに「まず」という表現が出てくるといったものが挙げられた。

4. 提案システムの設計・実装

4-1 提案システムの設計

4-1-1 形式的な誤り

● 文字数の指摘

本提案システムでは、送られてきた文章が指定文字数の 8 割以下であった場合、指定文字数を満たしていないと指摘することにした。

● 全角英数と半角カナ文字の指摘

本提案システムでは、半角カナ文字をレポートに不適切な文字とし指摘することにした。また、全角英数字についても、レポートにおいては半角英数字を推奨される [16] 場合が多いため、推奨項目として指摘することにした。

指摘の基準は、MeCab²⁾で形態素解析された単語を正規表現によりマッチさせ、全角英数字や半角カナ文字が混ざっていた場合とした。

● 1 字下げの指摘

本提案システムでは、段落ごとに 1 字下げが行われているか確認する。段落ごとに文章解析を行った際に、改行しかない段落は、そのまま改行の表示をする。次に、文頭が空白文字(1 字下げ)から始まり、その後に空白文字が連続しなかった場合、1 字下げがされているため、解析した文章をそのまま表示する。空白文字が連続した場合は、

2) オープンソースの形態素解析エンジン

正しく1字下げがされていないため、1字下げができていないことを指摘する。また、文頭に空白文字(1字下げ)がなかった場合でも、文末に句点がない場合、見出しや参考文献などの可能性が高いため、例外として指摘を行わないようにする。句点がきていた場合の多くは、1字下げがされていないため、1字下げができていないことを指摘する。

4-1-2 文法の誤り

● である体・ですます体の混合の指摘

本提案システムでは、である体・ですます体の混在があった場合、レポートを書くにあたって一般的に推奨されている「である体」^[17]に直すよう指摘する。そのため、本提案システムでは、「ですます体」で統一された文章の指摘も行うことにした。

● 体言止めの指摘

本提案システムでは、形態素解析により句点の前の単語が名詞であった場合、体言止めであると判定し指摘する。

● 話し言葉の指摘

本提案システムでは、話し言葉や幼稚表現など、レポートにふさわしくない単語があった場合、csvファイルに登録した単語を読み込み、正規表現によりマッチさせることにより指摘する。

● ら抜き言葉の指摘

本提案システムでは、ら抜き言葉を形態素解析に従い指摘する。ら抜き言葉は、基本的に上一段または下一段活用の動詞+接尾語の「れる」によって成り立つため、その法則に従って指摘する。

1. 活用形の確認

形態素解析により、活用形が上一段または下一段活用の動詞の場合、ら抜き言葉の

可能性があるフラグを立てる。

2. 接尾語「れる」があるかどうか

次の単語を見ていく時、1でら抜き言葉のフラグが立っていた場合、接尾語「れる」があるかどうかを確認する。接尾語「れる」がある場合、ら抜き言葉である指摘をする。

また、否定形の場合は、接尾語「れる」の判定部分で、助動詞・活用形「特殊・ナイ」及び「特殊・ヌ」の判定を行い当てはまった場合、ら抜き言葉である指摘をする。

● い抜き言葉の指摘

本提案システムでは、ら抜き言葉を形態素解析に従い指摘する。い抜き言葉は、基本的に自立動詞+非自立動詞「てる」及び「でる」の基本形で成り立つため、その法則に従って指摘する。

1. 品詞の確認

形態素解析により、品詞が自立動詞の場合、い抜き言葉の可能性のあるフラグを立てる。

2. 活用形及び単語の確認

次の単語を取り出した時、その単語が非自立動詞の基本形で且つ「(立)てる」「(読)ん」であるという単語であった場合、い抜き言葉の指摘を行う。

また、本提案システムでは、い抜き言葉の否定形の指摘は不可能である。い抜き言葉の否定形は、現在進行形か完了形かの違いで判定が異なってくるためである。例えば、「並べている」の否定形は「並べていない」であり、「並べてない」と書いた場合、い抜き言葉となる。しかし、「並べてある」の否定形の場合、「並べてない」となり、これは正しい文法である。この現在進行形か完了形かという判定は、形態素解析では不可能であるため、本提案システムでは未実装である。

● 単語の不統一の指摘

本提案システムでは、「こと」と「事」と

いった単語や「当たる」と「当る」といった送り仮名の不統一を指摘する。本研究では、単語の不統一は名詞、送り仮名の不統一は動詞で主に出現するものとしたため、それ以外の品詞での単語脳統一は指摘を行わない。

まず、形態素解析したファイルを読み込み重複単語の除外をした後、単語のソートを行う。単語のソートを行った後、それぞれの処理を以下の通りを行う。

1. 単語(名詞)の不統一

本研究では、単語の不統一を名詞に限定しているため、ソート後は名詞の抽出をし、前後の単語を比較する。この比較した単語の読みが同じだった場合、表層形(単語そのもの)が同じであるかどうか判定する。これが異なった場合、比較単語が双方とも漢字であるかどうかを判定する。これは、例えば「以降」と「移行」といったような全く違った意味であるが、読みが同じである単語の指摘を防ぐためである。双方とも漢字でなかった場合は、本提案システムでは単語の不統一と判定し、指摘を行う。

2. 送り仮名の不統一

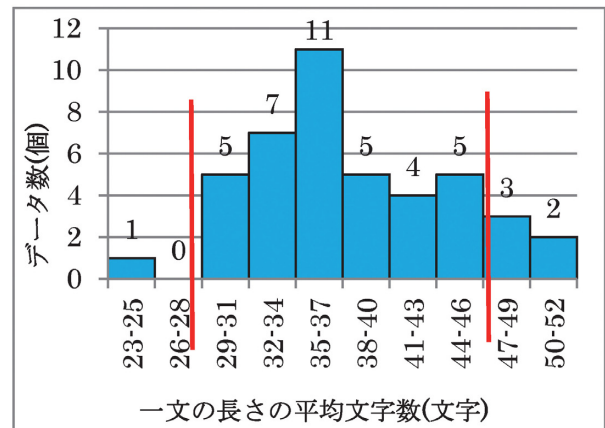
送り仮名の不統一は、多くの場合動詞において出現するため、ソート後は動詞の抽出をし、前後の単語を比較する。この比較した単語の読みが同じだった場合、単語の原形が同じであるかどうか判定する。これが異なる場合、本提案システムでは送り仮名の不統一と判定し、指摘を行う。

4-2 指摘指標の決定

指摘指標は、日本語表現 I のレポートを用いたヒストグラムによる分析で決定した。

● 一文の長さ(平均)

提案システムでは、1つの文章に対する一文の長さの平均についてグラフ6の結果から下限と上限の決定をおこなった。



グラフ6. 一文の長さの平均

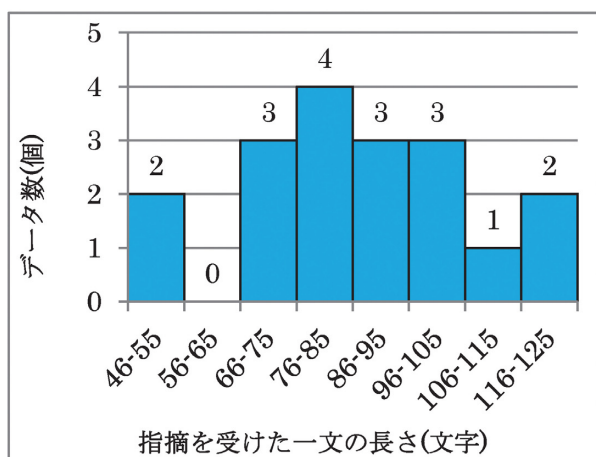
下限の指摘は、各一文が単文及び箇条書きであったり、word文章での体裁を整えるために不自然な改行が入ったりする可能性を示唆するために行う。決定方法としては、まず文献調査を行った。その結果、一文の文字数平均で、最も少ない文字数を挙げていたのが30文字から35文字^[18]であったことから一般的に30文字程度が限度であるとし、下限30文字を仮指標とした。次に、本分析に使用したレポートにおいてこの仮指標に該当した人数は、3人であった。このうち平均文字数が23文字であったレポートを分析すると、多くの単文が修正されていたことから本指摘は、正しいものとした。しかしながら、残り2人の平均文字数が29文字のレポートは、それぞれ評価がAとCであり、仮指標の数値があまり適切ではないと判断した。これらの結果から、本提案システムでは、一文の長さの平均下限は、28文字を本指標として決定した。

また、上限の指摘は、読み手にわかりやすい文章にするための重要な項目である。決定方法は、文献調査と本学教員へのヒアリング調査によるものとし、上限は、本学日本語表現の講義でも指導されている40文字を仮指標とした。次に、本分析に使用したレポートにおいてこの仮指標に該当した人数は、全体の約1/3である14人であった。更に、該当レポートの分析を行った結

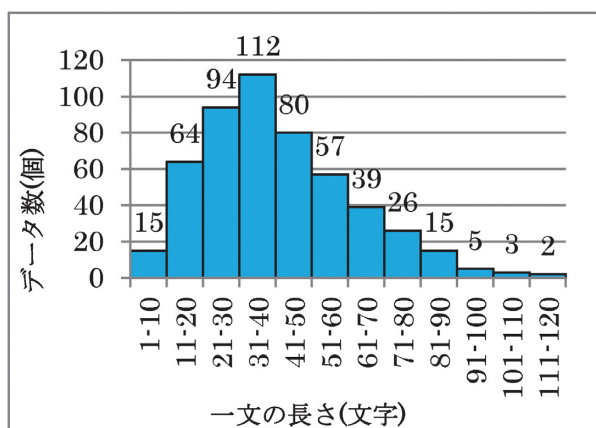
果、一文の平均文字数が41文字から45文字のレポート8部において、A及びBの評価を得ているレポートが3部あった。これは、良文のレポートのうち半数が指摘されることとなる。また、一文の平均文字数が41文字から45文字のレポート8部のうち、実際に一文が長いと指摘を受けているレポートは、1部のみであった。これらの結果から、仮指標の数値はあまり適切ではないと判断し、本提案システムでは、一文の長さの平均上限は、46文字を本指標として決定した。

● 一文の長さ

一文の平均文字数の下限と上限が決まったところで、次に一文ごとの長さを指摘する指標をグラフ7、8の結果から決定した。



グラフ7. 指摘を受けた一文の長さ



グラフ8. 一文の長さ

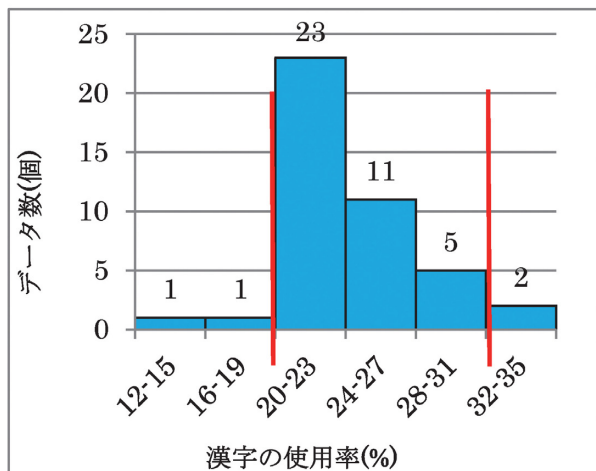
レポートの中で一文が長いと指摘を受けた一文を抽出した結果、18個のデータが得られた。そのデータから長さを分析すると最低が46文字、最高が118文字であった。一文の長さは、60文字を超えるとわかりにくくなる^[19]とされていることから、上限の仮指標を60文字とした。この仮指標を適用した場合、すべての一文のうち約2割、そして全レポートのうち約8割が指摘を受けることになった。更に、良文の分析を行った結果、65文字前後の一文も複数見受けられたため、これらの結果を踏まえ本提案システムでは、一文の長さを指摘する上限指標を65文字とした。

● 漢字の使用率

本提案システムでは、漢字使用率の指摘も行う。これは、漢字の使用率が少ない場合、ひらがなの羅列といったような読みづらい文章となっている可能性が高いためである。またその逆に、多すぎる場合も、過度な漢字変換をしている可能性があるため指摘を行うことにした。

漢字使用率の下限と上限は、グラフ9の結果から決定した。まず、漢字の使用率の下限について、文献では20%以下の文章は締まりがない文章とされている^[20]こと、また、レポート分析により良文の評価を得ている文章でも漢字使用率が20%であることから、20%を本指標とした。また、上限については、文字数と漢字使用率を同時にチェックするWEBライティング用ツール^[21]の基準を使用することにした。本ツールでは、漢字の使用率の基準を、新聞の社説(約50%)>マニュアル(約40%)>教科書(約30%)>雑誌(約30%)>WEB(約25%)であるとし、一般文章における判定基準を、20%以下は少ない、20%を超えて30%以下は適切、30%を超えると多いとしている。この

基準から、漢字使用率の上限は30%を本指標とした。

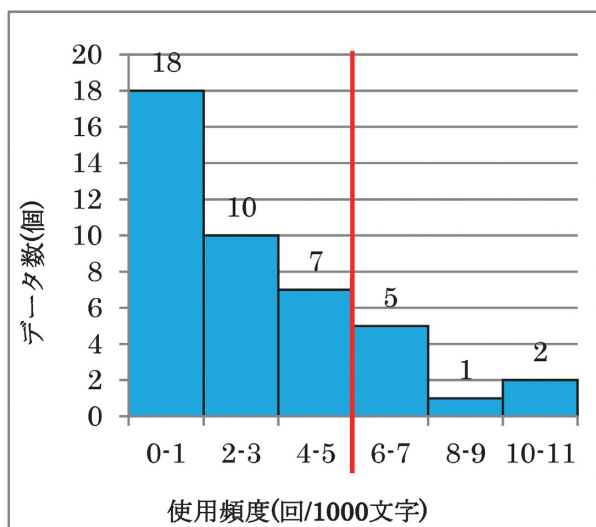


グラフ 9. 漢字の使用率

● 推定表現の使用頻度

本提案システムにおける推定表現の使用頻度の指摘は、グラフ 10 の結果から決定した。グラフにおける推定表現の使用回数は、1000 文字中出现する回数である。

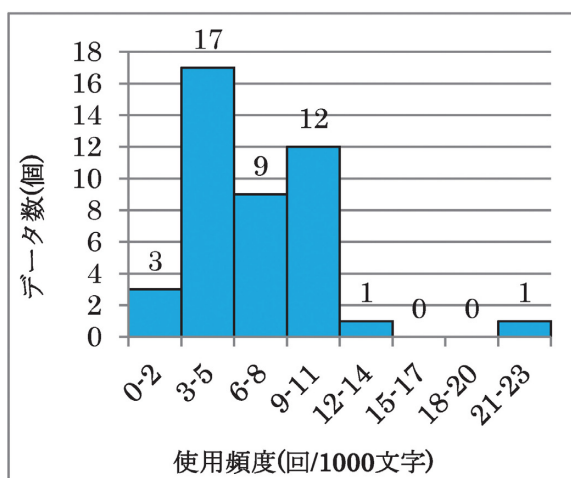
本分析で使用したレポートは、課題が自分の体験談を踏まえて述べたものであるため、「思う」「感じる」といった推定表現が通常より多いものと考えた。そのため、推定表現の使用回数の上限は、良文における最高使用回数の 5 回とした。



グラフ 10. 推定表現の使用頻度

● 指示代名詞の使用頻度

本提案システムにおける指示代名詞の使用頻度の指摘は、グラフ 11 の結果から決定した。指示代名詞の使用は、一文の中で 2 回以上あると好ましくないものとされている [22]。本提案システムでは、一文の長さの平均上限が 46 文字であることから、一文で 1 度ずつ指示代名詞が使われるとして、1000 文字中 12 回以上の指示代名詞が出現した場合、指摘するものとした。



グラフ 11. 指示代名詞の使用頻度

4-3 提案システムの実装

本提案システムは、学生の煩わしさを少しでも解消するため Web アプリケーションとした。提案システムは、図 1 のように Web を経由して文章を投稿すると、MeCab による形態素解析と一文の長さ、漢字の使用率といった文章や先ほど形態素解析を行った単語の解析を行う。そしてこれらの解析を基に、指摘画面の生成を行い、結果画面を返すようになっている。

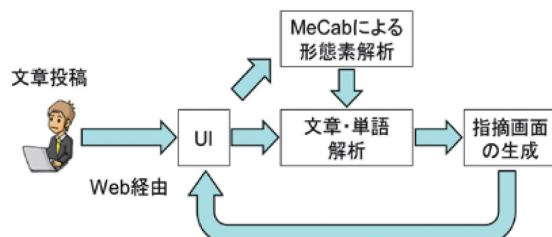


図 1. 提案システムの指摘画面の生成

5. 試作システムの評価・検証

本項目では、試作システムの指摘と教員による指摘の比較を行い、本試作システムの指摘の有効性と教員の負担軽減に対する有効性を検証した。

比較対象は、分析にも使用した日本語表現 I のレポート計 43 部である。試作システムの指摘率*は、下記の式(1)によって算出する。

$$\text{指摘率} = \frac{\text{本ツールによる指摘総数}}{\text{教員による指摘総数}} * 100$$

・・・(1)

5-1 形式的な誤り

文章量、氏名の記載不備、見出しの欠落などのレポートの形式について、本試作システムにおける指摘は、20.0%可能だった。詳細は、次の通りである。

- 文章量(指定の文字数)

本試作システムでは、指定の文字数の基準を 8 割以上とした。この基準における教員の指摘に対する指摘率は、100%であった。

また、本試作システムの指摘結果は、43 のサンプルのうち 23 のレポートにおいて指定文字数が足りなくなった。しかしながら、教員による添削で明確に文章量について指摘しているのは、2 部である。本項目に関しての考察は、2 つ挙げられる。1 つは、訂正箇所や削除箇所が多く、完成されたレポートではないためである。つまり、文章量以前の問題であり、これから修正が入ることを考慮し、現時点での文章量指摘は重要性が低いものとされていると考えた。もう 1 つは、明確に文章量についての指摘がされていないだけで、具体性をもっていない内容について更なる説明を促す指摘があるためである。この具体性をもっていない

内容に説明を付け足すことで、文章量が足りると判断され、指摘されていないものであると考えた。

- 課題内容に不適切な見出し

本項目は、文章内容に携わるものであるため、試作システムでの指摘は不可能であった。レポートの形式における誤りのうち、文章量以外の指摘は、すべてこれに当てはまる。

- 全角英数字・半角カタカナ

本検証で使用したレポートは、手書きの文章であるため、全角英数字及び半角カタカナの判別は、困難であった。ただし、原稿用紙上におけるアルファベットの書き方を誤っているレポートは多数見受けられた。

- 段落の一字下げ

提出レポートにおいて段落の 1 字下げは、正確にされていたため、試作システムでも指摘はなかった。本検証において手書きの文章は、段落の 1 字下げの意識が強いということが考察される。

5-2 文章ルール(文法)の誤り

ですます体・である体の混在、体言止めなどの基本的な文法の誤りについて、本試作システムにおける指摘は、98.2%可能だった。詳細は、次の通りである。

- ですます体・である体の混在

提出レポートにおいて、ですます体・である体の混在は見受けられなかった。また、ですます体の文体で書かれたレポートも見受けられなかった。

- 体言止め

提出レポートの体言止めについて、教員の指摘に対する指摘率は、100%であった。

- 話し言葉

話し言葉については、教員の指摘コメントを基にしたため、教員の指摘に対する指摘率は、100%であった。しかし、話し言葉のコーパスは、一般的にレポートに適していないとされる基本的な用語と本レポートにおける教員の指摘コメントを基にしているため、その指摘は完全とは言えない。

- ら抜き言葉

提出レポートのら抜き言葉について、教員の指摘に対する指摘率は、100%であった。

- い抜き言葉

提出レポートにおいて、い抜き言葉は見受けられなかった。しかしながら、試作システムは、い抜き言葉の否定形の指摘が不可能であることから、100%の指摘は難しいものであると考える。

- 単語の不一致

提出レポートにおいて、単語の不一致を指摘されているコメントは、数字が漢字とアラビア数字が混ざっていたレポートにしかされていなかった。そのため、本項目における教員の指摘に対する指摘率は、100%となった。しかしながら、本試作システムでの指摘の結果は、「こと」と「事」、「とき」と「時」、「できる」と「出来る」などの単語表記の不一致が24のレポートで見受けられるようになった。本検証における考察としては、本検証に使用したレポートは手書きの文章であるため、非常に漢字の誤字が目立っていた。そのため、誤字の修正の優先度が高く、単語の不一致に対する優先度は低いものなのではないかと考えた。

- その他

検証で使用したレポートの文章ルールにおける誤りのうち、上記の項目以外にカギ括弧の使い方の誤りが指摘されていた。カギ括弧の使い方は、文章内容によって変わってくるため、本試作システムでの指摘は不可能であった。

5-3 文章の質の低下要因

本試作システムでは、一文の長さ、主述のねじれ、曖昧表現の多用といった文章の質の低下要因の指摘が55.0%可能だった。詳細は、次の通りである。

- 一文の長さ(平均)

検証に使用したレポートでは、一文の長さの平均についての指摘はなかったため、有効性は検証不可能である。しかし、本試作システムにより、不自然な改行が入った文章や箇条書きが多い文章の数は、減るのではないかと予測する。また、分布図からも一文の長さの平均が悪文では2極化していることから、本項目は有効性があるものだと考えた。

- 一文の長さ

提出レポートの一文の長さについて、教員の指摘に対する指摘率は、9割弱であった。また、本試作システムの指摘結果は、すべてのレポートから約80箇所において一文が長いとなった。これは、教員による指摘の4倍以上を本試作システムにおいて指摘していることになる。一文の長さについては、諸説が多く存在するため、本研究でもまだ改善の必要がみられた。

- 述語のねじれ

主述のねじれは、文章内容に携わるものであるため、本試作システムにおける指摘はほぼ不可能であった。本試作システムで

は、「私の考えは～考える。」といった二重表現の指摘のみ可能である。また、主述のねじれを指摘された一文は同時に、一文の長さについても指摘されているため、本試作システムによる一文の長さの指摘により改善がみられる可能性があると考えた。

- 漢字の使用率

本検証に使用したレポートでは、漢字の使用率についての指摘はなかったため、有効性は検証不可能である。ただし、漢字使用率の分布を見ても、極端に漢字使用率が低い文章は低い評価となっているため、本項目は有効性があるものだと考えた。

- 推定表現の使用頻度

提出レポートの推定表現の使用頻度について、教員の指摘に対する指摘率は、100%であった。また、本試作システムの指摘結果は、提出レポート43部のうち8部が推定表現の使用頻度について指摘を受けているとなった。ただし、本検証に使用したレポートは、自分の体験談を踏まえて述べたものであるため、本レポートに対しては妥当であると考えた。

- 指示代名詞の使用頻度

本検証に使用したレポートでは、指示代名詞の使用頻度についての指摘はなかったため、有効性は検証不可能である。

これらの結果から、本試作システムは、教員による指摘総数のうち約5割を指摘したこととなった。これにより、本試作システムは教員の負担軽減が可能であり、本学学生の文章力向上に寄与するものであると考えた。

6. まとめ

本稿では、本学学生の日本語力及び提出レポートの問題点について挙げ、これによる本学教員の悩みと負担について述べた上で、本学学生の日本語文章力向上を本研究の目的とした。この目的を達成するために、本研究では2つの方法を選んだ。1つは、学生が自ら改善箇所を確認し、自発的に推敲をすることで、同様の問題を繰り返さないようにすることである。もう1つは、教員が従来、基本的な形式・文法の誤りといった箇所の指摘に割かれていた時間を軽減することである。

第2章では、既存類似システムの比較と有効性についての確認を行った。その結果、これら既存のシステムは、指摘精度が低いものか、指摘精度が高くても本学学生が利用するにあたって問題点があった。そこで、本学学生の日本語文章力向上のために、機械処理による文章の改善箇所を自動的に指摘する本学学生に向けた自発的に文章の推敲を促すためのシステムを提案した。また、提案システムを試作するにあたって、本学学生に向けたシステムとするために、文献調査及び教員へのヒアリング調査を通じて文法・形式的な誤りといった基本的な指摘項目を洗い出した。それと同時に、本学学生の提出レポートについて、文章の質低下要因を分析し明確化し、質低下の主要な原因となる指摘項目を挙げた。

第4章では、提案システムを実装するための処理内容について述べた。また、前章で挙げた質低下の主要な原因となる指摘項目の指摘指標を統計的手法によって分析を行い決定した。試作システムをWebアプリケーションの形で実装し、これらの指標をもとにシステムの指摘画面を生成し、結果をブラウザ上に表示する構造とした。

さいごに、既存システムの比較・評価・検証を行った。また、本試作システムの指摘と教員による指摘の比較によって、本試

作システムの指摘の有効性と教員の負担軽減に対する有効性を検証した。その結果、本試作システムは、教員による指摘総数のうち約5割が指摘可能であり、教員の負担を軽減することができるものと考えた。また、これにより少なくとも本学学生の文章力向上に寄与可能であることを確認した。

「日本語文章表現授業に対する作文支援システム導入手法の開発」『日本教育工学会研究報告集』10(1), pp.93-100, 日本教育工学会

[5] 吉岡 友治(2013)『いい文章には型がある』PHP 研究所.

[6] 野口 悠紀雄(2002)『「超」文章法』中央公論新社.

[7] 山内 志郎(2001)『ぎりぎり合格への論文マニュアル』平凡社.

[8] 木下是雄(1994)『レポートの組み立て方』筑摩書房.

[9] 村岡 貴子・因 京子・仁科 喜久子(2013)『論文作成のための文章力向上プログラム—アカデミック・ライティングの核心をつかむ』大阪大学出版会

[10] 斎藤 美奈子(2002)『文章読本さん江』筑摩書房.

[11] 久恒 啓一(2008)『図で考えれば文章がうまくなる』PHP 研究所.

[12] 三省堂編修所(1995)『何でもわかることばの知識百科』三省堂.

[13] 山内 志郎(2001)『ぎりぎり合格への論文マニュアル』平凡社.

[14] 高山瞭「文章力アップの方法／無料講座&原稿添削リライト案内」,
<<http://bunsyounet.net/kouza1/09syoryaku.html>> 2015年1月7日最終アクセス

[15] 石黒 圭著(2008)『文章は接続詞で決まる』光文社.

[16] 関西学院大学総合政策学部(2012)『基礎演習ハンドブック—さあ、大学での学びをはじめよう！(改訂新版)』関西学院大学出版会.

[17] 金沢大学 共通教育機構「レポート作成

参考文献

[1] 下條 正純(2010)「論述式答案に見る日本人大学生の日本語諸問題」『佐賀大学留学生センター紀要』9, pp.47-59, 佐賀大学

[2] 『東京読売新聞』記事「大学の實力 教育力向上の取り組み(上)」(2011年7月6日)

[3] 日本語検定委員会『「大学生の日本語力」第1期・第2期調査レポート』(2012年3月1日)

[4] 山口 昌也, 北村 雅則, 棚橋 尚子(2010)

の手引き」,
<<http://www.kanazawa-u.ac.jp/faculty/ki-ko/students/report.pdf>>2015年1月7日最終アクセス

[18] 辰濃 和男(1994)『文章の書き方』岩波書店.

[19] 樺島 忠夫(1980)『文章構成法』講談社.

[20] 『東京読売新聞』記事「編集手帳」(2010年11月27日)

[21] 文字数と漢字使用率を同時にチェック
| WEBライティング用ツール
<http://www.gal-moji.com/moji/kanji_ritsu.asp>2015年1月7日最終アクセス

[22] 又平 恵美子, 竹内 純人, 大野 博之
(2010)「文章作成支援ツールによる日本語文章力育成」『論文誌 ICT活用教育方法研究』13(1), pp.16-20, 私立大学情報教育協会