

ICTを活用した日本語文章力育成への取り組み：問題点と展望

Promoting Academic Writing Education through Information & Communication Technology

- Problems and Prospects -

稲積宏誠† Hiroshige Inazumi
大野博之‡ Hiroyuki Oono

1. まえがき

近年、母語としての日本語表現力の低下、特に正確で簡潔な表現による論理的な文章作成能力の低下は、教育現場を中心に重大な問題と位置づけられている。ここで言う日本語表現力とは、事実を客観的に示し、明確に主張を述べることができること、また、複数の解釈を与えず、冗長な表現を含まない文章を作成することのできる表現力とする。すなわち、レポート、論文、マニュアル等の技術文書などの作成に求められる素養に相当する。

また、文章力は、伝えたいことを正確に表現するという意味において、伝えるべき情報がどれほど整理されているか、すなわち論理的であるかが問われる。

著者らは、論理的思考力の育成を意識しつつも、まず、論理的な展開となるような文章表現上のシンタックスを満足した表現力を狭義の論理的な文章と位置づけ、その育成を研究目的と位置づけた。

そこで、著者らは、文章表現上のシンタックスを効率的に学習できる環境の構築を目的とし、文章表現法を、単に知識として覚えるのではなく、文章上の問題点に自ら気づき、考え、文章校正を行うというプロセスを自然に身につけられる環境の構築を目指した取組を続けてきた。[1]

本稿では、この取組のなかで、表記や文構造の問題点を捉え正しい文章表現に改善していくための「校正・推敲支援ツール」を取り上げる。これは、「文章のもつ曖昧さの解明」や「難解な表現の解釈」に注目してきた自然言語処理技術と ICT 全般を、文章力育成のための技術に転嫁する取組と位置づけ、人工知能、特に自然言語処理技術と、情報システム開発の立場から、正確で簡潔な表現で論理的な文章作成を支援するものである。

また、本稿では特に母語話者としての学生、特に初学者を対象とした従来の取組を発展させ、非母語話者を対象とした取組への移行について検討することを目的とする。まず、現在開発を続けている校正・推敲ツールの特徴を明らかにし、非母語話者の典型的な誤用をどのように解釈、処理することができるかについて考察していく。

2. 校正推敲支援ツールの概要

学習者が書いた文章を学習者自身で校正・推敲していくための支援環境を提供するのが本ツールの目的である。これは、誤り箇所の指摘や注意すべき箇所の指摘、文章改善のヒントなどを提供するものである。すなわち、校正・推敲ルールを生成し、任意の文章において校正・推敲対象項目の検出を可能とすることによって、文章表現として正誤情報を正しく判定することのできる環境を生み出すことで

ある。

ツール設計にあたっては、汎用的なプラットフォーム化を実現するために、校正・推敲を行う機能は、plug-in と呼ばれるソフトウェアに対する機能拡張の仕組みにより実現されている。これによって、必要に応じて機能の拡張・修正が可能となる。さらに、plug-in の仕様を公開することで、自然言語処理技術の容易な活用も支援している。自然言語処理の結果は、どの plug-in から利用可能な汎用的なモジュールとして準備されている。すなわち、フリーウェアとして利用可能な形態素解析ツールである Mecab や JUMAN、係り受け解析ツールである Cabocha や KNP を対象とし、複数の文からなる文書を 1 文ごとに切り分け、形態素解析・係り受け解析を行い、形態素情報・文節情報・係り受け情報を構造的に変換する汎用モジュールである。

さらに、この汎用モジュールは、ダイナミックリンクライブラリ(DLL)形式とし、ソフトウェアから動的に利用することを可能とした。これによって、各種 plug-in からこのようなモジュールを動的に参照することで、文書の解析結果を容易に利用することができる。

文章を執筆する際には、いくつかの定型的な校正ルールに従う必要がある。まず、これらの校正ルールは、ルールベースとして整理しなければならない。さらに、執筆対象となる領域に依存した校正ルールも存在することから、それらに柔軟に対応する必要がある。現在のところ、本稿の対象とする技術文書に加えて、どの分野にも共通して利用可能と思われるものを中心に、文書作成に関する書籍 [2]-[4]などを参考にして校正ルールを定めている。

それぞれの校正ルールは、ルールベースに対して、日本語解析処理による文節の表記、係り受けの結果、および形態素の表記や品詞などの情報を加味することで実装されている[5]。

(1) 基本機能

本ツールは、図 1 に示すとおり、4 つの機能に対応するウィンドウから構成される。

- 文書の編集を行うテキストエディタ領域(左上)：既存文書の読み込み、新規作成、修正等を行う。
- 文書全体の属性情報領域(文字数や漢字含有率、異なり語数など)の表示領域(右上)：文章に関する基礎情報として保持される。
- 一文ごとの校正・推敲対象の指摘表示領域(右下)：一文の中のどの箇所に誤り候補が存在するか、誤りの種別と位置を示す。
- 指摘項目ごとの詳細表示領域(左下)：誤り候補についての詳細情報を表示する。必要に応じて、修正候補についても表示する。

(2) 応用機能

より高度な校正・推敲機能については、現在開発中のものが多い。ただし、特に、長文での曖昧さを強く意識して、

†青山学院大学, Aoyama Gakuin University

‡東京医療保健大学, Tokyo Healthcare University

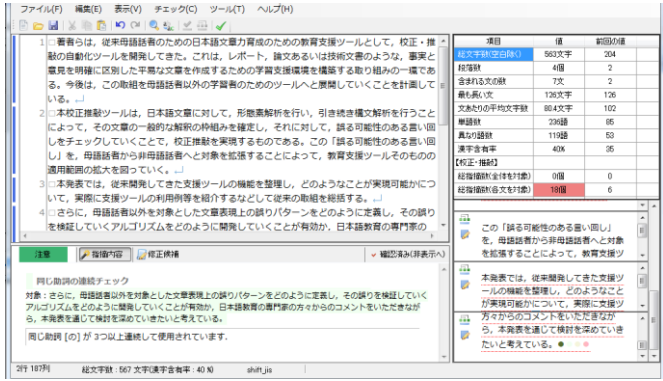


図1 校正・推敲支援ツールの概要

係り受け関係に基づいた文節単位での構造を視覚化することによって、何が解釈の違いを導くのかについて検討することを可能としている。

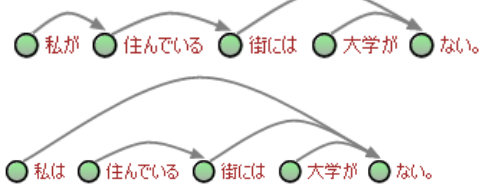
次項で示す非母語話者を対象とした取組では、まずこの係り受け関係を分析するための機能を利用した支援環境を検討することとする。

3. 非母語話者への展開

非母語話者の誤用分析については多くの調査報告があるが、基本的にそれぞれの母語のもつ言語的な特徴との対比から生まれるものと、日本語独特の表現上の特徴から生まれるものが混在しているようである。ここでは、中国人を対象とした調査結果[6]を参考にして、いくつかの誤用表現に注目し、本ツールの機能〜どのような取組が可能であるかを検討する。

(1) 連体節における主語

連体節における主語として助詞「が」で受けるべきところを「は」を使用してしまう場合が良く見られる。しかし、これを本支援ツールで解析すると、



となり、連体節における主語として機能しなくなることが判定されることになる。さらに、ツールに組み込まれている節推定機能を用いると、



という表示され、連体節が生成されていることを示すことが可能である。ただし、この場合には、誤用と判定された後の解説機能は存在するものの、「は」の利用が誤用であるとの指摘を行うためには、そのルール化が必要となる。

(2) 助詞「の」の過剰使用

助詞の使いかたのなかで、格助詞「の」の過剰使用は誤用の典型例とされる。この場合には、「かわいい女の子が好きだ」についての下記の解析結果からわかるとおり、形容詞に助詞が接続して連体修飾となるため、この点が誤

り指摘のルール化のポイントであることがわかる。

* 0 1D 0/1 1.907573 かわいい 形容詞自立**形容詞-I段基本形 の 助詞連体化***のノ 0	* 0 1D 0/0 1.469137 かわいい 形容詞自立**形容詞-I段基本形 の 助詞連体化***のノ 0
* 1 2D 0/1 0.000000 女の子 名詞一般***女の子オナノコ は 助詞係助詞***はハフ 0	* 2 -1D 0/1 0.000000 好き 名詞形容動詞語幹***好きスギ だ 助動詞***特殊-I基本形だだ 0
* 2 -1D 0/1 0.000000 好き 名詞形容動詞語幹***好きスギ だ 助動詞***特殊-I基本形だだ 0	EOS

一方、「きれいな／の）人は得をする」のように、形容動詞に助詞の「の」をつけてしまうような誤りも、容易に判定できると考えられる。

* 0 1D 0/1 1.441783 きれいな 名詞形容動詞語幹***きれいな の 助詞連体化***のノ 0	* 0 1D 0/1 1.109438 きれいな 名詞形容動詞語幹***きれいな の 助詞連体化***のノ 0
* 1 3D 0/1 0.000000 人 名詞一般***人トトト 0 は 助詞係助詞***はハフ 0	* 1 3D 0/1 0.000000 人 名詞一般***人トトト 0 は 助詞係助詞***はハフ 0
* 2 3D 0/1 0.000000 得を 名詞一般***得トクトク を 助詞格助詞一般***をヲ 0	* 2 3D 0/1 0.000000 得を 名詞一般***得トクトク を 助詞格助詞一般***をヲ 0
* 3 -1D 0/0 0.000000 する 動詞自立**サ変-I基本形する	* 3 -1D 0/0 0.000000 する 動詞自立**サ変-I基本形する

ところが、「たくさん／の）客に感謝する」のように、助詞の接続関係により、形容動詞と副詞のいずれかに解釈されることがある。このように、品詞と品詞間の接続関係に加え、語彙のもつ詳細な情報を加味しなければならないことがわかる。

* 0 1D 0/1 1.320442 たくさん 名詞形容動詞語幹***たくさん の 助詞連体化***のノ 0	* 0 1D 0/1 1.643258 たくさん 名詞副詞可能***たくさん の 助詞連体化***のノ C
* 1 2D 0/1 0.000000 客 名詞一般***客キヤクキヤク に 助詞格助詞一般***にニ 0	* 1 2D 0/1 0.000000 客 名詞一般***客キヤクキヤク C に 助詞格助詞一般***にニ 0
* 2 -1D 1/1 0.000000 感謝する 名詞サ変接続***感謝カンシャ する 動詞自立**サ変-I基本形する	* 2 -1D 1/1 0.000000 感謝する 名詞サ変接続***感謝カンシャ する 動詞自立**サ変-I基本形する C

4. 今後の取組

本稿では、日本語文章力育成に向けた著者らの取組みである校正推敲支援ツールを紹介し、その展開の一つとして、非母語話者支援についての展望の一端を紹介した。

開発の指針としては、誤用表現の指摘精度の違いをもとにして次の3点から取り組む予定である。

1. 誤用表現を特定するルール化を実現させ、任意の文章中の誤用表現の自動チェックを実現させる。
2. 誤用表現候補となる表現候補すべてを無条件に指摘し、誤用か否かの判定条件を提示し、それをもとに各自判定するための補助情報を提供する。
3. 誤用表現を対象とした正誤問題の解説資料として、日本語文章の解析結果としてのツールからの出力情報を活用する。

これらの取組については、次稿以降順次発表していく予定である。

参考文献

- [1] 稲積宏誠他, "ICT を活用した日本語文章力育成への取り組み", 情報処理学会 コンピュータと教育研究会, Vol.2011-CE-109 (2011).
- [2] 木下是雄: 理科系の作文技術. 中公出版(1981)
- [3] 中島利勝, 塚本真也: 知的な科学・技術文書の書き方. コロナ社 (1996)
- [4] 野口悠雄: 「超」文章法 - 伝えたいことをどう書くか-. 中公出版(2002)
- [5] 大野博之, 稲積宏誠: 技術文章の校正・推敲支援ツールにおける機能拡張容易性の向上, 電子情報通信学会技術研究報告, 教育工学研究会, pp.31-36.(2008)
- [6] 張麟声 "日本語教育のための誤用分析—中国語話者の母語干渉20例", フリーエースネットワーク (2001).