

令和元年6月10日現在

機関番号：32402

研究種目：基盤研究(B) (一般)

研究期間：2015～2018

課題番号：15H03219

研究課題名(和文) 多義語の意味の自動特定機能を組み入れたやさしい日本語による読解支援環境の構築

研究課題名(英文) Developing a web-based system for simplifying and rewriting Japanese polysemous words and phrases

研究代表者

川村 よし子 (Kawamura, Yoshiko)

東京国際大学・言語コミュニケーション学部・教授

研究者番号：40214704

交付決定額(研究期間全体)：(直接経費) 12,300,000円

研究成果の概要(和文)：本研究の目的は、やさしい日本語書き換えシステム「やさしくな－れ」を多義語の意味の特定も可能なものに改良することにある。本研究では、多義語の意味の決定パターンの抽出を行い、それをもとに「意味特定アプリ」を開発し、「やさしくな－れ」に組み入れた。「意味特定アプリ」は、これまでの形態素解析による分析に加えて、慣用句等への対応を可能にするため、連語を扱える仕組みを取り入れた。最新版の「やさしくな－れ」では、入力文と書き換え文が併記される。また、適切な書き換えが難しい語や意味の特定が難しい語については、書き換えは行わず、バルーンを用いて、意味をやさしい日本語で示す。研究成果はweb上で無償公開している。

研究成果の学術的意義や社会的意義

共生社会実現のためには、日本語弱者に対する支援が不可欠である。日本語弱者にとって、漢字熟語が多く含まれ、カタカナ語も多用された日本語の文章は難解であり、情報の提供方法の改善に加えて、読解支援のための対応が求められている。本研究は、これまで開発を進めてきたやさしい日本語への自動書き換えシステム『やさしくな－れ』を改良し、多義語の意味の特定が可能なものに整備した。特に、漢字熟語が多いサ変名詞や形容動詞等について、前後関係から意味を特定しやさしく書き換えて表示する仕組みを整えた。また、入力文と書き換え文の双方に自動で漢字の読みを表示できるようにした。システムはweb上で公開している。

研究成果の概要(英文)：This research was undertaken to improve and expand the functions of the web based rewriting system known as Yasashikuna-re (a system for rewriting Japanese words into simplified words). Specifically, the research aimed at the inclusion of polysemic words (words with multiple meanings) and word phrases (such as idioms) to broaden the range of words that can be simplified. In order to include polysemic words and word phrases, the authors developed a semantic meaning identifier application based on previous research on semantic patterns found in these words, and then integrated this application into the system. The new version of Yasashikuna-re shows a screen with both the simplified results and the original sentence. In the case of words that are not amenable to simplification for various reasons, these words are marked with notes in a balloon format on the screen giving the definition of the word in easy Japanese. The new system is available with open access on the web.

研究分野：日本語教育

キーワード：日本語教育 読解支援 共生社会 やさしい日本語 多義語 コロケーション 慣用表現 Web環境

様式 C - 19、F - 19 - 1、Z - 19、CK - 19 (共通)

1. 研究開始当初の背景

本研究の社会的背景として、共生社会実現のために日本語弱者に対する支援が不可欠であるという事情がある。日本語弱者にとって、漢字熟語が多く含まれ、カタカナ語も多用された日本語の文章は難解であり、情報の提供方法の改善に加えて、読解支援のための対応が求められている。これまで、研究代表者は、日本語学習者への読解支援システムの提供を行うとともに、外国人介護福祉士候補者のための教材作成プロジェクトチームを主導してきた。その活動を通して、日本が「共生社会」をめざすためには、わかりやすい日本語で書かれた文章による情報発信を行う必要性を痛感し、やさしい日本語への自動書き換えシステムの開発を行ってきた。

研究の学術的背景としては、研究代表者らが、1999年から、Web上で公開してきた日本語読解学習の支援サイト『リーディング・チュウ太』(<http://language.tiu.ac.jp/>)がある。このサイトはWeb上の学習支援環境として先駆的役割を果たしてきたが、提供している学習支援ツールの一つに、入力文中のすべての単語の難易度を自動で判定する「語彙チェッカー」がある。これは、文章に含まれる単語の難易度を「旧日本語能力試験」の出題基準をもとに判定し、級ごとに色分けして表示するツールである。「語彙チェッカー」は、やさしい日本語の文章を書く際の有用なツールとして、『やさしい日本語』の作成ルール(弘前大学)「外国人にやさしい日本語表現の手引き 2006」(埼玉県総合政策部国際課)を始め、様々な公共機関の広報紙等で紹介され、各方面で活用されている。その後、文の難易度の判定システムとして、「帯」(Sato, *et al.* 2008)や、「日本語リーダビリティ測定」(柴崎 2009)などが公開されたが、いずれも読み手として日本語母語話者を想定しているため、日本語学習者のためには「語彙チェッカー」に代わるものはなかった。そこで、研究代表者らは、平成 21-23 年度基盤研究(B)において、個々の単語の難易度および一文ごとの長さを基準にして文章の難易度を自動判定するシステムを開発し(川村 2011)、Web上で公開(<http://basil.is.konan-u.ac.jp/chuta/level/>)した。

さらに、前研究課題である平成 24-26 年度基盤研究(B)では、日本語教師のやさしい文への書き換え術を「集合知」として収集し、文章に含まれる難易度の高い単語を自動でやさしい単語に書き換えるシステムの開発を行い、『チュウ太のやさしくな-れ』(以下『やさしくな-れ』)としてWeb上での公開を開始した。このシステムの開発過程において、多義語の意味を特定する必要性が明らかになったため、今回の研究を開始することとなった。

2. 研究の目的

本研究は、多義語の意味を特定し、自動でやさしい日本語に書き換えることのできる読解支援システムの構築を目指している。日常よく使われる基本語の中には、意味範囲が広く、慣用表現等にも用いられ、多義語になるものが数多くあるが、難易度の高い単語にも多義語が多数存在する。多義語の意味の特定は文脈に依存しているため、単純に自動でやさしい日本語に書き換えることはできない。そこで、既存の『やさしくな-れ』では、多義語については書き換えを行わず、複数の意味を併記して示すという形をとっていた。ところが、「日本語弱者」にとって多義語の意味の特定は難しい。そのため、難易度の高い多義語についても、文中での意味を特定したうえで、やさしい日本語に書き換える必要性は高い。そこで、本研究において、特に漢字熟語から構成されている難易度の高いサ変動詞・形容動詞等について意味決定のルールを抽出し、多義語の意味の特定が可能な意味特定アプリをシステムに組み入れることにした。

3. 研究の方法

本研究では、使用頻度が高くかつ難易度の高い多義語の収集および多義語の意味の決定パタ

ーンの抽出を行うとともに、そこから得た知見を基に、多義語の意味を自動で特定するアプリケーション（以下「意味特定アプリ」）を開発する。この「意味特定アプリ」は、評価実験を経て完成させ、入力文を自動でやさしい日本語に書き換えるシステム『やさしくな－れ』に組み入れて公開する。研究の流れは次の通りである。

- (1) 多義語を例文とともに収集し、やさしい日本語による書き換えリストを作成する
- (2) 例文を分析し、多義語の意味の決定パターンを抽出する
- (3) 意味の決定パターンによって多義語を分類する
- (4) 書き換えの有効性に関して学習者による評価実験を行う
- (5) 評価実験をもとに書き換えシステムの改良を行う
- (6) 多義語の意味の特定が可能な「意味特定アプリ」を開発する
- (7) 日本語教師の協力を得て「意味特定アプリ」の運用実験を行う
- (8) 運用実験の結果をもとに「意味特定アプリ」の改良を行う
- (9) 日本語学習者を対象として、多義語対応の『やさしくな－れ』の運用実験を行う
- (10) 多義語対応の『やさしくな－れ』を、共生のための読解支援環境として無償公開する

以上によって、Web 上に「日本語弱者のための読解支援環境」を提供することが可能になる。運用実験は、日本語の母語話者および世界各国の日本語学習者の協力を得て行い、その結果をもとに改良を加えて、実用的価値の高い日本語読解支援環境として提供する。

4．研究成果

(1) 難易度の高い多義語の選定

難易度の高い単語の選定は、UniDic の語彙素約 16 万語をもとにした。このうち、初級（旧日本語能力試験の出題基準をもとにした 3 級と 4 級）の語以外を対象とした。語の選定基準としては、新聞コーパスにおいて出現頻度の高いものを抽出し、形容動詞（形状詞）とサ変動詞（名詞・サ変可能）を中心に多義語リストを作成した。

(2) コロケーションに配慮した書き換えリストの作成

やさしい日本語への書き換えリストは、単語レベルではなく、前後の語句にまで配慮し、難易度の高い表現とやさしい日本語での表現を併記したリストを作成した。リストの作成にあたっては、形態素 N-gram コーパスおよび新聞コーパスを活用し、多義語の各意味を特定しうる条件を抽出した。また、慣用表現や連語等、コロケーションによって意味の変わる語に対しての配慮も行った。

(3) 「意味特定アプリ」の開発

上記のコロケーションに配慮した書き換えリストをもとに、各活用形にも対応しうる「意味特定アプリ」を開発した。また、サ変動詞+名詞によって構成される複合語に関しては、コーパスをもとにして前項となる動詞の活用形を特定する仕組みも整えた。「意味特定アプリ」は、『やさしくな－れ』に組み入れる形で開発した。なお、適切な書き換えが難しい語や上述した対応では意味の特定が難しい語に関しては、書き換えは行わず、バルーンを用いて、やさしい日本語で意味を説明することにした。

(4) システムの有効性に対する運用実験をもとにした改良

今回開発したシステムの有効性を検証するために、日本語学習者を対象に運用実験を行った。運用実験の結果をうけて、出力画面では、入力された本文と書き換え文を並記することにした。また、入力された本文・書き換え文・バルーン内の説明文に含まれるすべての漢字に自動で読みを表示する等の改良も行った。

(5) システムの公開

研究の成果物であるシステムを Web (<http://yasashiku.chuta.jp/>) 上で無償公開することにより、本研究の目的である多義語の意味の自動特定機能を組み入れたやさしい日本語による読解支援環境の提供が可能になった。このシステムは、日本語弱者だけでなく、やさしい日本語を媒介に情報を提供する必要がある機関等での活用も可能である。なお、本システムでは、書き換えリストをもとにすべての書き換えを自動で行っているため、やさしい日本語への書き換えが適切ではない場合もある。そのため、本研究終了後も書き換えリストの整備は継続していく予定である。

5. 主な発表論文等

[雑誌論文](計 12 件)

- (1) 八木真生・川村よし子 (2019) 「非漢字圏中上級学習者は読解中の辞書使用・不使用の判断をどのように行っているか 辞書引き行動の観察及びインタビューの結果から」 『日本語教育方法研究会誌』 査読無, Vol.25 No.2, 70-71
- (2) 川村よし子 (2018) 「日本語学習者のための web 辞書」 『ヨーロッパ日本語教育』 査読有, Vol.22, 66-72
- (3) 八木真生・川村よし子 (2018) 「『やさしくな-れ』は学習者の読解に役立つのか 日本語学習者を対象にした運用実験の結果から」 『日本語教育方法研究会誌』 査読無, Vol.24 No.2, 100-101
- (4) 北村達也・川村よし子 (2017) 「新聞記事における慣用表現の出現頻度調査」 『甲南大学紀要知能情報学編』 査読無, Vol.10 No.1, 25-34
- (5) 金庭久美子・川村よし子・山森理恵・小林秀和 (2017) 「日本語学習者に役に立つ語は何か 「やさしい日本語」書き換えリストの分析を通して」 『日本語教育方法研究会誌』 査読無, Vol.23 No.2, 88-89
- (6) 金庭久美子・川村よし子・橋本直幸 (2016) 「作文支援ツール『文体チェッカー』の開発と評価」 『日本語教育方法研究会誌』 査読無, Vol.23 No.1, 66-67
- (7) 川村よし子 (2016) 「逆引き辞書の作成を可能にする対訳辞書編集システムの構築」 『ヨーロッパ日本語教育』 査読有, Vol.20, 285-290
- (8) 北村達也・川村よし子 (2016) 「新聞記事における慣用表現の出現頻度の調査」 『日本語教育方法研究会誌』 査読無, Vol.22 No.3, 32-33
- (9) 篠崎大司・松下達彦・川村よし子 (2016) 「上級日本語文字・語彙 e ラーニング教材の開発」 台湾日本語言語文藝研究学会 『日本言語文藝研究』 査読有, 第 6 号, 90-108
- (10) 北村達也, 孝橋一希 (2015) 「文難易度の調整を支援するシステム『かぶとエディタ』の連語対応」 『日本語教育方法研究会誌』 査読無, Vol.22 No.2, 20-21
- (11) 八木真生・川村よし子 (2015) 「サ変名詞+接尾辞『者』をやさしい日本語へ書き換える形態素 N-gram を利用した自動処理の検証」 『日本語教育方法研究会誌』 査読無, Vol.22 No.2, 62-63
- (12) 川村よし子 (2015) 「日本語教師の集合知を活用したやさしい日本語書き換えシステムの構築」 『ヨーロッパ日本語教育』 査読有, Vol.19, 131-136

[学会発表](計 15 件)

- (1) 金庭久美子・川村よし子 (2019) 「慣用表現に対応したやさしい日本語書き換えシステム」 日本語教育学会支部集会

- (2) 川村よし子 (2108)「AI 時代到来に備えた日本語教師の役割」台湾日本語教育学会 2018 国際シンポジウム
- (3) 八木真生・川村よし子 (2018)「辞書を使用した文章理解過程の分析 非漢字圏中級日本語学習者を対象として」CAJLE2018
- (4) 川村よし子 (2018)「読解教育および学習支援のための機能表現抽出ツールの開発と運用実験」AJE2018
- (5) 川村よし子 (2018)「外国人児童・生徒ための日本語教育」言語文化教育学会シンポジウム『内なる国際化の一例：外国人児童・生徒の問題』
- (6) 川村よし子 (2018)「IT 時代に対応した日本語教師の役割」韓国日語教育学会学術大会
- (7) 川村よし子 (2018)「読解支援システムへの AI 活用の可能性と留意点」AATJ2018
- (8) 金庭久美子・川村よし子・山森理恵 (2017)「連語に配慮したやさしい日本語書き換えシステムの構築」韓国日本語学会国際学術大会
- (9) 川村よし子 (2016)「東南アジア圏学習者のための日本語辞書編集システムの開発」ICJLE2016
- (10) 砂川有里子・李在鎬・長谷部陽一郎・川村よし子 (2015)「日本語学習辞書開発プロジェクトについて」日本語教育学会秋季大会
- (11) Tatsuya Kitamura (2015) An editing system with a word-level identification function for adjustment sentence difficulty, Construction of Digital Resources for Learning Japanese
- (12) 金庭久美子・川村よし子 (2015)「レベル差のあるクラスのための教材開発 『やさしくなーれ』を利用した教材作成」韓国日本語学会国際学術大会
- (13) 川村よし子・川村ヒサオ (2015)「多言語対応『チュウ太の Web 辞書』の最新版の活用」CASTEL/J 2015
- (14) 川村よし子 (2015)「やさしい日本語書き換えシステムの提案機能とその活用」CASTEL/J 2015
- (15) 北村達也・田中豪 (2015)「教材作成を支援する ICT ツールの紹介」日本語教育学会実践研究フォーラム

〔図書〕(計 4 件)

- (1) 川村よし子 総監修 (2017)『ポータブル日インドネシア英・インドネシア日英辞典』三修社
- (2) 川村よし子 総監修 (2016)『ポータブル日マレー英・マレー日英辞典』三修社
- (3) 川村よし子 総監修 (2015)『ポータブル日タイ英・タイ日英辞典』三修社
- (4) 川村よし子 総監修 (2015)『ポータブル日ベトナム英・ベトナム日英辞典』三修社

〔その他〕

ホームページ等

- (1) チュウ太のやさしくなーれ <http://yasashiku.chuta.jp/>
- (2) チュウ太の Web 辞書 <http://chuta.jp/>
- (3) 表現文型抽出器 http://eins.overworks.jp/phrase_analyzer/
- (4) リーディング・チュウ太 <http://language.tiu.ac.jp/>
- (5) マレー語版チュウ太 <http://malay-chuta.overworks.jp/>

6. 研究組織

(1) 研究分担者

研究分担者氏名：北村 達也

ローマ字氏名：Kitamura Tatsuya

所属研究機関名：甲南大学

部局名：知能情報学部

職名：教授

研究者番号（8桁）：60293594

研究分担者氏名：金庭 久美子

ローマ字氏名：Kaneniwa Kumiko

所属研究機関名：立教大学

部局名：日本語教育センター

職名：特任准教授

研究者番号（8桁）：60733772

研究分担者氏名：前田 ジョイス（2016年度まで）

ローマ字氏名：Maeda Joyce

所属研究機関名：東京国際大学

部局名：言語コミュニケーション学部

職名：教授

研究者番号（8桁）：90275800

研究分担者氏名：八木 真生（2016年度から）

ローマ字氏名：Yagi Maki

所属研究機関名：東京外国語大学

部局名：世界言語社会教育センター

職名：特任助教

研究者番号（8桁）：90782904

(2) 研究協力者

研究協力者氏名：小林 秀和

ローマ字氏名：Kobayashi Hidekazu

研究協力者氏名：保原 麗

ローマ字氏名：Hobara Rei

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。