

科学研究費助成事業 研究成果報告書

令和 元年 6 月 24 日現在

機関番号：37402

研究種目：基盤研究(C) (一般)

研究期間：2016～2018

課題番号：16K00500

研究課題名(和文)類似文検索を用いた日本語教育用コンテンツ開発環境の整備

研究課題名(英文)A Contents Development Environment for Japanese Education using Similar Syntax Search Function

研究代表者

堤 豊(Tsutsumi, Yutaka)

熊本学園大学・商学部・教授

研究者番号：80227446

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：本研究では、日本語教育教師のための学習コンテンツ作成支援システムを開発した。このシステムの最大の特徴は、MS-Word上で完結したシステムということである。従って、コンピュータに不慣れな教師であっても普段利用しているMS-Wordのメニュー画面からKWIC(Key Word In Context)や単語の難易度の情報を知ることができる。

また、本システムでは類似文検索機能を提供している。本機能は、例示した文と構文的に似た文をデータベース中から検索し、表示する機能であり、似たような構文をたくさん探し例示することで学習者の理解を深めることができる。

研究成果の学術的意義や社会的意義

本研究では、MS-Wordをベースにした日本語教育支援システムの開発を行った。研究成果は国際交流基金主催のマレーシア日本語教育学会等で発表し、好評を得た。高価で特殊なソフトウェアを必要とせず、普段使い慣れているMS-Word上で語学教育で必要とするほとんどすべての機能が利用できることを示したことは意義深いと考えられる。また、類似文検索機能についても教育上の効果が高いことはわかっているが、これを簡単に利用できることを示せたことは当初の目的を達成できたと考えられる。

研究成果の概要(英文)：This research project aimed at providing a content development system for JSL teachers. The most important function of this system is that it is a MS-Word integrated environment. It means JSL teachers do not need any other software.

This system also allows teachers to use a similar syntax search function. It helps teachers to show same syntax sentences in their class, or to make their own test.

研究分野：情報工学

キーワード：類似文検索 日本語教育

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

1. 研究開始当初の背景

昨今、高等教育において「アクティブラーニング」が注目を浴びている。語学教育においては、自ら学ぶ姿勢がないと身につかないという意味で、親和性の高い教育方法だと考えられる。また、語学教育においては教材提示や発音の学習においてコンピュータでの自習学習は歴史のある教育方法である。一方で、アクティブラーニングで効果があるといわれている、協調学習については多数のシステムが開発されているにも関わらず、あまり実践的に効果が上がっていない。この理由として、これらのシステムが独自に開発されたシステムであり、教員や学生が使い慣れた環境とは違うため、勉強しなおさないといけないことや、教員が新たな機能を付け加えることができないことが挙げられる。

この問題点を解決するために、筆者らは、従来研究で、一貫して教材作成を簡単にするためのソフトウェアの開発を行ってきた。例えば、WEB 上で問題を作成できるようなツールを提供し、ワープロ感覚で問題が作成できるツールの開発などが挙げられる。これらの研究では、英語を母語としない学生(EFL)の英語教育を対象に研究してきた。

また、申請者らは九州大学牛島研究室で開発された日本語文章校正ツール「推敲」のソースコードの保守を行っている。「推敲」では、日本語文章のパターンを正規表現で切り分けることで高速に助詞、助動詞の分類を行うことができる。この技術を使って、日本語文の類似検索システムを教材ツールとして実現した。本研究では、「推敲」の技術と申請者らが今までに培ってきた英語教育のための教材作成ツールの技術を基にして、アクティブラーニング実現のための協調学習環境上で簡単に利用できる例文検索機能を実装する。

2. 研究の目的

研究の目的は、日本語教育において教師が簡単に学習コンテンツを作成するために必要な例文検索機能を、特別なソフトウェアを導入することなく教員・学生が普段利用している MS-Word 上に構築することである。筆者らが従来開発してきた類似文検索システムと統合し教材となる文を簡単にインターネット上の膨大な文書中から検索・提示し、文・単語単位に注釈やマーカーをつけグループ間で情報共有しながら学習が進められる、使いやすいツールとして提供する。また、並行して従来から研究している英語教材作成支援ツールの開発も行い、相互の研究成果を活かせるようにする。

3. 研究の方法

本研究を達成するに当たり、開発すべき要素技術は、Word 中の日本語文章に対して(1)形態素解析し内部的に DB として記憶し、(2)各文および単語にタグ付けしノートや辞書引き、蛍光ペンのマークなどの結果を保存・参照できるようにする、(3) VSTO アドインソフトにより学生間でのグループワークを可能にする技術、(4)教材サーバーにてユーザとグループおよび課題について管理する技術を開発する。また、(5)筆者らが従来開発してきた類似文検索システムおよびテキスト解析ツールを統合することが必要となる。上記(1)(2)については、筆者が主に研究を続けてきたテーマである。(3)(4)については、長年マルチメディアによる教材作成システムを手掛けてきた熊本県立大学名誉教授松野了二氏を研究協力者として招へいし、2名の研究体制で開発した。

4. 研究成果

4-1. はじめに

「普段利用している Microsoft Word (以下、Word と呼ぶ) 内で教材作成等に関するすべての作業を行えばどれだけ楽になるであろう」

すべての作業を Word 内で行うことはさすがに無理であるが、日本語教育に関連する一連の作業に必要な機能をできるだけ多く Word のメニューとして提供し、教員の手間を少しでも省けるようにすることが本研究の目的である。例えば、読解力教育を例にとると、教材用資料収集、テキスト解析(難易度や用例検索など)、協調学習、理解力確認問題作成、評価作業等が必要であろう。いうまでもなく、教育に利用可能なアプリケーションソフトウェア(以下、アプリと呼ぶ)は現在でも数多く存在する。例えば、テキスト解析ツールを用いた読解力の教材作成や Moodle などの eラーニングプラットフォームを用いたオンライン学習コースの作成などに多くの日本語教育者がコンピュータを利用している。しかし、教育に利用可能なアプリの現状にはいくつかの問題点もある。例えば、Moodle の利用にはサーバーが必要であり手軽に利用できない教員も多い。また、Word で教材を作成中にテキスト解析を行いたい場合、テキスト解析アプリを起動し、何らかの方法で解析したいテキストをそのアプリに渡し、解析結果を再び Word に戻し作業を続けなければならない。これは手間もかかるし、操作方法もアプリごとに異なることが多く煩雑な作業となる。この点、一連の作業を Word 内で完結できれば複数のアプリを切り替える必要もなく、操作方法にも一貫性があるものを提供できる。本研究では、このような問題点を解消すべく、インターネット上にある膨大なテキストから指定した用例を含む文を検索し取り込む機能、テキスト解析機能、協調学習支援データベース作成機能、など前述した一連の作業を Word のメニューとして提供し、Word 内で作業を完結可能とすることを目指して

いる。本稿では日本語教育支援ツール（以下、WIDEJL¹と呼ぶ）の概要について述べた後、本ツールの中核であるテキスト解析部を中心に紹介する。

4-2 . 日本語教育支援ツールの概要

我々はこれまでいくつかの教育支援ツールの研究・開発を行ってきた。この中で WIDEJL に直接結びついているものを文献[5]、[6]に示す。これまでの研究では、「授業に使用する教材は学生のレベルに合わせたものを手作りすることが重要であり、手作り教材を手軽に作成できるツールを提供したい」ということを基本理念としている。なお、ここでいう「手軽に」は、「1」ウェブと表計算の基本操作が可能な程度の知識があれば、2」 Moodle など管理者が必要なサーバーを使える環境になくても」を意味する。

本ツールもこの理念に基づいたものであり、日々の授業中において、学生の理解度をチェックしてみたいと思ったときなどに教材を手軽に手作りできるツールとなることを目指している。以下に WIDEJL の概要について述べる。

4-2.1 WIDEJL の概要

図 1 に WIDEJL の概要図を示す。WIDEJL ではテキストの解析に MeCab²を利用している。また、

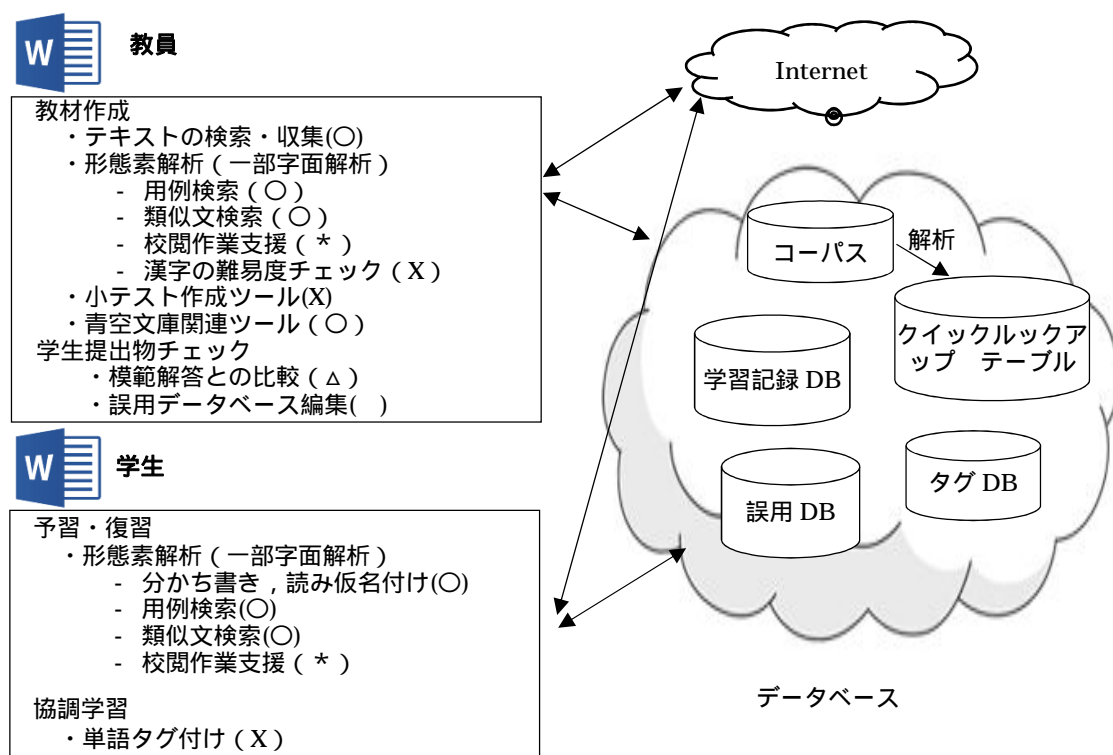


図.1 WIDEJL の概

青空文庫³をコーパスとして利用可能なように青空文庫用のユーティリティプログラムも用意した。以下に概要について述べる。紙面の関係上すべてを説明することはできないが、(○)がついている項目は開発済み、() はほぼ開発済み、(X) は開発中または設計中の項目である。なお、(*) がついている項目のメニューの一部は Word が本来提供しているメニューと重複しているが、操作の一貫性を考慮し敢えてメニューに加えた。また、図 1 において教員用と学生用に分けているが、これは使用対象者を想定して分類したものでありメニュー上で区別しているわけではない。

4-2.2 WIDEJL のツール

1) 教材の検索、収集ツール・・・我々が開発した WIDEJL と連携可能な Web ブラウザを利用して、インターネット上にある膨大なデータから指定した用例を検索し、Word 上に取り込むことが可能である。

2) テキスト解析ツール・・・Word 上に取り込んだ(入力した)テキストを解析する。本機能の核には Mecab を利用して形態素解析を行い実現している(一部には字面解析も利用)。テキス

¹ この略称は“MS Word Integrated Development Environment for Creating Japanese Language Learning Materials”からとった。

² 京都大学情報学研究所-日本電信電話株式会社コミュニケーション科学基礎研究所共同研究ユニットプロジェクトを通じて開発されたオープンソース形態素解析エンジン

³ 著作権が消滅した作品や著者が許諾した作品のテキストを公開しているインターネット上の電子図書館で収録作品数は 14287 (2017.7.23 現在)

ト解析機能は一般的なテキスト解析アプリにみられる、語句の頻度表、KWIC コンコーダンス等に加え、類似構文検索機能を備えている。

図2にWIDEJLのテキスト解析のメニューを示す。図から分かるように、WIDEJLでは品詞情報などMecabが提供する形態素情報を利用した検索も可能である。

3) データベースツール・・・上記の検索結果や解析結果をデータベースに登録し、必要に応じて再利用可能である。また、青空文庫からダウンロードしたテキストを形態素解析しデータベース化する機能も備えており、さらに学習者の誤用データベース編集ツールも開発中である。



図2. WIDEJLのテキスト解析メニュー

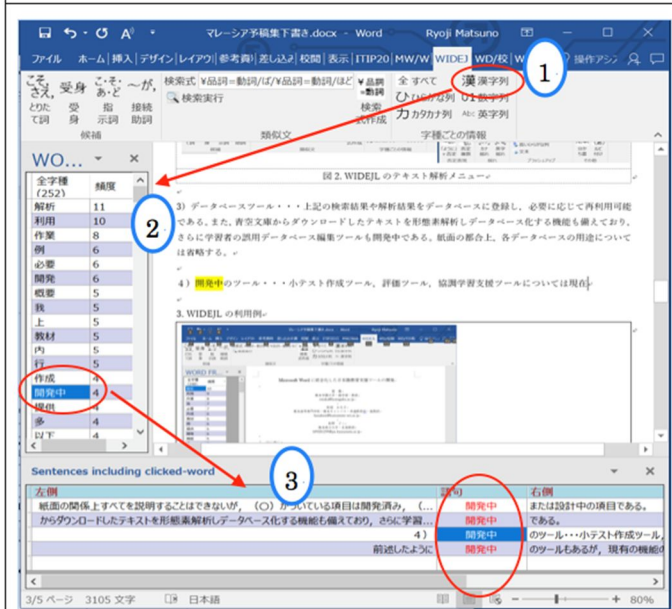


図3. WIDEJL 使用例—漢字頻度表

【WIDEJL 使用例—漢字頻度表】

- ①の「漢字列」メニューを選択すると②の頻度表ウィンドウが表示される。
- ②の頻度表の任意の漢字を選択すると③のKWIC ウィンドウが表示されるとともに解析対象のWord 文書中の対応箇所が黄色で色付けされる。
- ③のKWIC ウィンドウの任意の1行を選択すると、Word 文書中の対応箇所が選択状態となる。

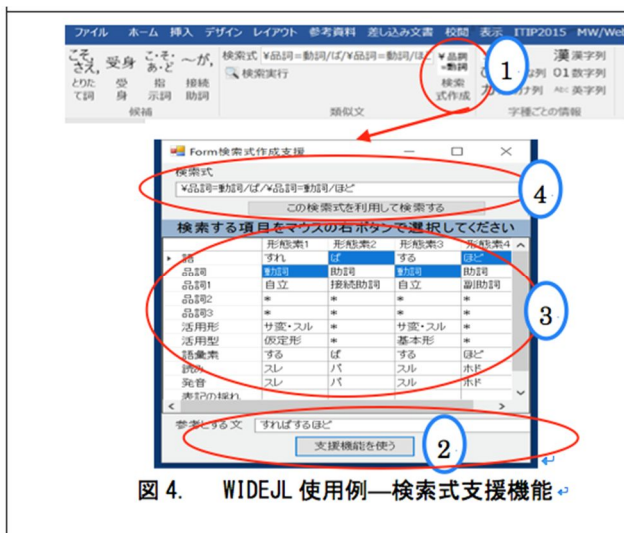


図4. WIDEJL 使用例—検索式支援機能

【WIDEJL 使用例—検索式支援機能】

- Word 文書中の任意の文字列を選択後、①の「検索式作成」メニューを選択すると「検索式作成支援ウィンドウ」が表示される。
- ②の参考とする文を編集後「支援機能を使う」ボタンをクリックすると③のように形態素解析結果が表示される。
- 検索項目を選択すると④の検索式ボックスに検索式がリアルタイムで反映される。
- ④の「この検索式を利用して検索する」ボタンをクリックするとWord 文書を検索し、その結果がKWIC ウィンドウに表示される。

各データベースの用途については省略する。

4) 開発中のツール・・・小テスト作成ツール、評価ツール、協調学習支援ツールについては現在開発中である。それぞれの内容については省略する。

4-3. WIDEJL の使用例

図3は下書き中の本論文を解析し、漢字列の頻度表を表示したスクリーンショットである。また、図4は類似文の検索式作成支援機能を使用中の画面のスクリーンショットである。どちらも図の右側に説明を付加しているので参照していただきたい。

4-4 . おわりに

WIDEJL のテキスト解析関連のメニューの多くは 1990 年前後に九州大学の牛島研究室で開発された日本語文章推敲支援ツール『推敲』⁴をベースとしている。『推敲』は科学技術論文の推敲支援が主目的であったが日本語教育にも適用可能な部分が多い。WIDEJL では MeCab を用いたことにより、品詞等を用いた検索も可能となりフレキシブルな機能を提供できるようになった。なお、4-2 . 1 でも述べたとおり、メニューのいくつかは Word 本来のメニューとしても提供されているが、WIDEJL では操作及び表示の一貫性を考慮して別途提供した。また、WIDEJL は我々がこれまで進めてきた英語教育ツール MiWiT の手法を日本語教育に適用したものである。本研究グループには日本語教育の専門家がいらないのが弱点であり、本報告は専門家からみた必要な機能等についてご意見をいただけることを期待している。最後に、WIDEJL は今回発表した中核となるテキスト解析部は使用可能であるが、小テスト作成ツールや協調学習支援ツールなど残りの部分の完成を急ぎ、近い将来総合的な日本語教育支援ツールとして提供したい。

5 . 主な発表論文等

〔学会発表〕(計 3 件)

Yutaka Tsutsumi, Kanako Matsuno, and Ryoji Matsuno, Developing MS Word Add-in Tools for Japanese As a Second Language Education, 5th ICRTTEL 2018 - International Conference on Research in Teaching, Education & Learning, Singapore, Nov. 2018.

Yutaka Tsutsumi, Kanako Matsuo, and Ryoji Matsuno, Microsoft Word Integrated Japanese Education Assistant Tool, THE 14TH INTERNATIONAL CONFERENCE ON JAPANESE LANGUAGE EDUCATION IN MALAYSIA(マラヤ大学予備教育部日本留学特別コース(AAJ)、マレーシア日本語教師会(MAJLIS)、国際交流基金クアラルンプール日本文化センター共催), Kuala Lumpur, Malaysia, Oct. 2017.

Yutaka Tsutsumi, Kanako Matsuo, and Ryoji Matsuno, MS Word Integrated Development Environment for Creating Japanese Language Learning Materials, AETC2016 (The Asian Education Technology Conference), Singapore, Nov. 2016.

6 . 研究組織

(1)研究分担者 なし

(2)研究協力者

研究協力者氏名：松野了二

ローマ字氏名：Ryoji Matsuno

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。

⁴ 『推敲』は当時のオペレーティングシステム MS-DOS 上で動作する「字面解析」のシステムである。我々は 2002 年に牛島教授の依頼により同じ字面解析を用いて Windows 上に移植した。