

例文集からの日本語教育練習問題の柔軟な自動生成システム

Flexible Automatic Generation System of Japanese Language Exercises from Example Sentence Database

北中 佑樹^{*1}, 竹内 和広^{*2}, 西村 由美^{*3}
Yuki KITANAKA^{*1}, Kazuhiro TAKEUCHI^{*2}, Yumi NISHIMURA^{*3}

^{*1}大阪電気通信大学総合情報学部 (非常勤講師)

^{*1} Faculty of Information Science and Arts, Osaka Electro-Communication University

^{*2}大阪電気通信大学情報通信工学部

^{*2} Faculty of Information and Communication Engineering, Osaka Electro-Communication University

^{*3}関西学院大学

^{*3}Kwansei Gakuin University

Email: vana@takelab.jp

あらまし: 本研究では, 自然言語処理モジュールを用いた日本語教育練習問題の自動生成システムを提案する. 日本語教育は, 多様な分野で行われており, それぞれに対応する教材の作成が求められている. 本システムの利用により, 日本語教師は複数の形式の練習問題を容易に作成することが可能となる. 多様な分野に柔軟に対応するために, 教師が作成した例文を活用した練習問題生成の枠組みを述べ, 我々が実装した試作システムの紹介を行う.

キーワード: 日本語学習, 語彙・単語学習, 文法教育, LMS

1. はじめに

1980年代から日本語学習者の多様性が注目されるようになり, 学習者の様々なニーズに対応すべく教材が作成されてきた. 近年では医療福祉分野に従事する外国人人材が急増した際に, 教材・学習支援の不足が指摘された. 今後も, 様々な分野で活動する学習者に対する日本語教育の充実が課題となるであろう.

現在, 日本語学習者が活動する分野は多岐にわたり, 日本語学習の期間や, 学習の支援に関して学習者の置かれる状況は様々である.

このような状況に対応するために, 次の2点を満たす枠組みが必要であると考えた.

①多くの既存の初級教科書に用いられている文型学習に準じている.

②学習者の現実に即した語彙や表現を柔軟に取り入れられる.

上記をふまえ, 本研究は, 授業の準備や教室活動といった教育現場において作成される例文を用いることにより, 教師の負担を増すことなく, 自動的に練習問題を作成するシステムを提案する.

2. システムの特徴

本システムは, 既存の教材に不足する部分を補うための教師の負担を軽減すること, 学習者の現実に即した語彙や表現を学習する素材を提供することを旨とする. 主な対象者は初級レベルの学習者である. 学習者は既習の言語知識を有機的に関連させながら, 各自の現実に必要な語彙や表現を拡充しつつ, 実際の運用へと転移, 発展させていく. その転移, 発展の段階を促進する練習に適した問題集は不足しており, その補完は教師の個別の工夫・努力, または学習者の現場での習得にゆだねられている.

既に, 教師の授業準備や, 学習者の復習や独習を支援するシステムは多数提供されている. 例えば, 広く使われている教科書に準拠し, その補習教材として問題集を提供するものは, 配布プリントとして活用したり, 学習者のペースに合わせて復習を促すことができる. また, 例文を多く提示し, 詳細な説明を述べるタイプのもは, 独習を可能にし, 授業で理解不足の部分を補うことができる. これらは容

易に利用でき, 利便性が高いものの, 学習者の現実に即した問題を加えるといった発展性は持たない. さらに, 練習問題を自由に作成するために, 教師が問題の形式や内容を選び, Webで公開できるソフトも提供されている¹. しかし, これは紙媒体ではなく, オンラインでの問題作成を支援することを目的としており, 問題と選択肢を作るなどの, 問題の素材を作成する過程を容易にするものではない.

本システムの特徴は, 授業の準備, あるいは学習の過程で作成される現実に即した例文そのものを, 練習問題のリソースと考えていることである. 登録された例文をリソースとして練習問題を自動的に作成するため, 教師は練習問題として「問い」と「答え」あるいは「選択肢」を用意することなく, 問題の形式と学習項目を選ぶだけで, 練習問題が作成できる.

3. 本システムの構成

本システムは問題集生成システムと, そのリソースとなる例文集データベース (以下では例文集 DB と称す) から構成される. 学習者と教師 (以下では併せて利用者と称す) は Web 端末を通して問題集生成システムを利用することが可能である.

例文集 DB への例文登録は登録者 (教師) が例文登録システムへログインして行う. この際, 例文そのものに加えて情報付与が行われる. 登録者は, それが用いられる場面や機能といった情報を同時に登録する. 例文登録システムは, 登録された例文に対し, 文型や語彙情報の付与を系統的に支援する. 本システムの概要を図1に示す.

3.1 例文に対する情報付与

例文のデータベース化に向けて, 計算機支援による情報付与だけではなく, 対話の文脈に関する情報を与えることを検討した. 日本語教育における例文の「文脈」の重要性は, 度々指摘されており, 例えば川口 (2002) ⁽¹⁾は「誰が/誰に向かって/何のために」ある表現が行われるのかを記述した上で指導

¹ Hot Potatoes (<http://hotpot.uvic.ca>)

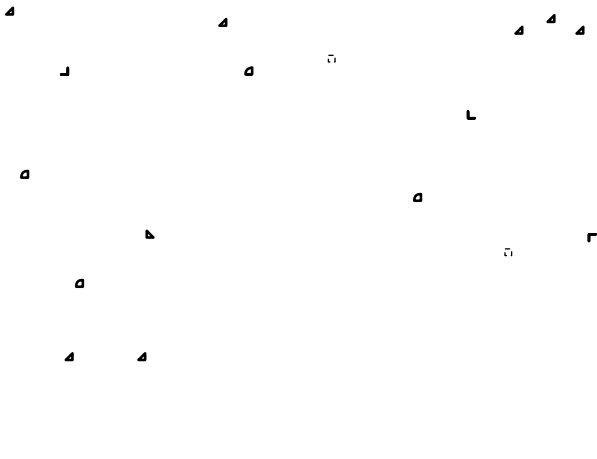


図1 問題集自動生成システムの概要

する、「文脈化」が必要であると述べている。システムによる日本語学習支援においても、練習問題作成のリソースとなり得る実用的な対話例を蓄積していくためには、文脈を示す情報を整理の基準として用いることが必要であると考え、「場面」と「機能」情報を付与した。

3.2 システムによる情報付与

例文をデータベースに登録する際に、より柔軟な問題集生成のために、計算機による情報付与支援を行う。

日本語教育において文法を学ぶこととは、文型と呼ばれる「ある機能を持つ表現」と、それに接続する「活用語の形」の組合せパターンを学習することである。例えば、「書いてください」という表現を学ぶ際には、「V てください」が学習文型として示され、V の部分を「書いて(て)」「見(て)」と入れ換えることによって文型表現を練習する。このため、ここでは、入力された例文がどの文型に該当するかを計算機により判定し、情報の付与を行なっている。

本システムで用いられている文型の一部を表1に示す。これらの文型は日本語教育で一般的に用いられている教科書⁽²⁾における記述方法を参考にした。文型には学習順序ごとに、1課から50課までの順序付けがされている。

表1 文型の例

文型	課
NはNです	1
これはNです	2
NをVます	6
AいNです	8

文型中にはNやVといったメタ記号が含まれており、Nは「名詞」、Vは「動詞」、Aは「い形容詞」などのように、対応する文構成要素で置き換えることができる。そのため、入力された例文と文型とを直接比較するのではなく、入力された例文に含まれる語がどの要素に該当するかを判別する必要がある。例えば、文型『NをVます』であれば、「パンを食べます」などの文が文型『NをVます』であると判断される。また、1つの例文に複数の文型が含まれることもある。

4. 日本語練習問題の自動生成

例文集DBを用いて、日本語練習問題の自動生成

を行う。学習の目的に合わせて、本システムでは以下に挙げる(a)から(f)の種類の問題生成を行うことができる。

(a) 漢字語彙：空欄に漢字の読み方を直接入力する。旧日本語能力試験⁽³⁾の1級から3級のレベル別に問題を設定することも可能である。

(b) 品詞別語彙：主に名詞・助詞・副詞といった活用変化のない語彙の知識を問うことを想定している。語形変換を問わない語彙問題として、動詞や形容詞を選択することもできる。利用者は問題文下部に示される選択群から語を一つ選び解答を入力する。

(c) 文型：選択群から動詞や形容詞を一つ選び、語形を変換させた上で入力する。適切な語彙の選択と文型に接続する語の変換を複合的に問うため、文の理解が前提となる。

(d) 文の接続：単文では、行動と対象物といった語と語の関係を、複文では原因と結果といった前件と後件の関係を問う。

(e) 文の構造：文を組み立てる能力を問う。意味と形式の関係の理解を促し、文レベルでの産出を目指す練習である。

(f) 対話：対話レベルでの意味理解に焦点を絞った練習である。第一発話に対して適切な応答が選べるかを問う。対話の場面を指定した練習も可能である。各種類の問題の具体的な例と回答の形式を表2に示す。

本稿で紹介した日本語練習問題の自動生成システムの試作は iitai (<http://takelab.org/iitai/>) にて稼働している。システムを利用する際は、生成したい問題の種類を選択し、問題数や問題の種類ごとのオプションを選択することで問題の生成が行える。

表2 問題の例と形式

種類	問題文の例	形式
(a)	夜中()に足がつかまりました 解答(よなか)	読み方入力
(b)	少し食欲()出ましたか 〈選択群:を, に, が〉	選択
(c)	熱もすっかり()ましたね 〈選択群:覚める 下がる 眠れる〉 (解答:下がり)	選択+語形の変形
(d)	〈前半文〉まだ眠いので 〈後半文〉起きたくないです	前半と後半の結びつけ
(e)	ご希望, 他, 何, ありませんか, は, に, か	並べ替え
(f)	〈第一発話〉今日は何を着ますか 〈応答〉ピンクのセーターがいいです	対話の結びつけ

5. おわりに

本稿では、柔軟な日本語教育練習問題の自動生成を目的として、例文への人手と計算機による情報付与を検討し、システムを試作した。場面や機能といった情報を与えることで、多様な分野に対応した例文の選択を可能とした。また、汎用的な自然言語処理モジュールを利用して、語彙情報や文型情報を機械的に付与することによって例文登録コストの軽減を図った。

参考文献

- (1) 川口義一:「文脈化」による応用日本語研究 文法項目の提出順再考, 早稲田日本語研究第11号, pp.57-63 (2003)
- (2) 財団法人海外技術者研修協会編:「みんなの日本語初級I, II本冊」, スリーエーネットワーク (2003)
- (3) 国際交流基金, 財団法人日本国際教育協会編:「日本語能力試験出題基準 改訂版」, 凡人社 (2002)