

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 31 日現在

機関番号：12608

研究種目：挑戦的萌芽研究

研究期間：2015～2016

課題番号：15K12166

研究課題名(和文)インターネット上の資源を活用した英語語彙問題の自動作成

研究課題名(英文) Automatic generation of English vocabulary questions using Internet text resources

研究代表者

徳永 健伸 (Tokunaga, Takenobu)

東京工業大学・情報理工学院・教授

研究者番号：20197875

交付決定額(研究期間全体)：(直接経費) 1,900,000円

研究成果の概要(和文)：本研究では、インターネット上のテキストを利用することによって、英語語彙問題を自動作成する新しい手法を提案し、その実用性について評価をおこなった。より多様で、時世に即した問題を作成するために、コンピュータによって試験問題を自動作成する研究がおこなわれているが、従来の研究が、あらかじめ人手で作成された辞書などの資源を使って問題を作成していたのに対して、本研究では、インターネット上のテキストを利用することによって、より新鮮で多様な問題を作成する新しい手法を提案し、作成した問題を実際に英語学習者に解かせることにより、問題の質が人間が作成した問題に比べて遜色がないことを確認した。

研究成果の概要(英文)：We proposed a novel method for automatically generating English vocabulary tests. The English vocabulary question is a multiple-choice question consisting of a target word, a reading passage, a correct answer and distractors. Given a target word, we generate a reading passage from Web texts retrieved from the Internet, and then employ that reading passage and the WordNet lexical dictionary for generating question options. Two kinds of evaluation were conducted considering two aspects: (1) measuring English learners' proficiency and (2) their similarity to the human-made questions. The first evaluation is based on the responses from English learners obtained through administering the machine-generated and human-made questions to them, and the second is based on the subjective judgement by English teachers. Both evaluations showed that the machine-generated questions had comparable quality with the human-made questions.

研究分野：計算言語学，自然言語処理

キーワード：自然言語処理 英語語彙問題 自動問題作成 言語学習

1. 研究開始当初の背景

グローバル人材育成の重要性が高まっている中、TOEICやTOEFLに代表される英語能力試験が活用されている。これらの試験の参考書は多く出版されているが、これらはみな人手で作られており、その問題作成には多大な労力を要する。一方で、インターネットユーザの約29%は英語の使用者であるという統計があり、これは英語テキストがインターネット上に大量に存在することを意味している。より多様で、時世に即した問題を作成するために、コンピュータによる言語処理技術を用いてこのような英語能力試験の問題を自動作成しようという研究、特にテキスト中で指定された語(句)と同じ意味を表わす選択肢を選ぶ形式で語彙能力を問う語義選択問題を作成する研究がおこなわれてきた。しかし、これらはいずれも人手で用意された資源を使って問題を作成している。我々はこれまでにインターネット上のテキストを利用して、既存の英語辞書に新しい情報を追加する研究をおこない成果をあげてきた。この成果を活かし、我々はインターネット上の膨大な英語テキストを利用して語彙問題を自動作成することを着想した。

2. 研究の目的

本研究では、インターネット上のテキストを利用することによって、英語能力試験の問題を自動作成する新しい手法を提案し、その実用性について評価をおこなうことを目的とする。具体的には以下の2つの研究項目を設定する。

【問題の自動作成】語の意味を選択肢から選ばせる語義選択問題では、まず、問題とする語がテキスト中でどの意味(語義)で使われているかを正確に同定する必要がある。これを実用上十分な精度で実現できる手法を開発する。さらに、同定した語義に基づいて、正解の選択肢と不正解の選択肢を作成する。最終的には与えられた語について、その語を含むテキストをインターネットから取得し、それに基づいて選択肢を生成し、問題を作成するプロトタイプを構築する。

【自動作成した問題の評価】従来は、自動作成した問題を英語教師によって主観的に評価する評価法が中心であった。この

主観評価に加え、本研究では英語学習者に自動作成した問題を解かせ、同一学習者の人手で作成した問題に対する解答と比較することにより、その実用性の評価をおこなう。

従来の研究は既存の辞書などを利用して問題を作成したのに対し、本研究ではインターネット上のテキストを利用する点に特色がある。これにより、最新のテキストにおける語の意味を問う問題が作成できる。また、従来の研究では、選択肢が一語に限定されているものがほとんどである。しかし、語の意味を複数の語で言い換えることも可能であり、このような複数語の選択肢を含む問題の自動作成はこれまで試みられたことがなかった。この研究の成果は、英語教師の問題作成の支援、また、英語学習者が自習時に自動作成された問題を解くことによって、多様な問題に触れる機会を与えるなどの効果が期待できる。

3. 研究の方法

以下の手順で研究を進める。

(1) 語義選択問題の分析

本研究で対象とする語義選択問題について、入試で作成された既存の問題を英語試験の参考書などから収集し、「大学英語教育学会基本語リスト」に基づくJACET8000英単語を参考にして、出題される語彙の範囲、難易度、正解と不正解の選択肢の関連性などについて分析をおこない、問題がどのような観点を考慮して作成されているかを明らかにする。収集する問題は500問から1,000問程度を予定しているが、必要に応じて増減することも検討する。特にコーパス言語学の知識を応用しながら分析検討をする。

(2) 指定した語義(語)を含む文脈検索

上述のとおり指定した語義の意味で語が使われている文脈(テキスト)をインターネット上から検索する処理は本研究の中核部分である。評価のための対象語として、参考書から100語程度を選択し、各語の語義を英語語彙WordNetを参考にランダムに選び、その語義で語が使われている文脈をインターネット上から検索する。評価用の100語の選択については(1)の分析結果から得られた知見をふまえて、問題作成の観点が偏らないように配慮する。検索はGoogle, Yahoo, Bingな

どの一般の検索エンジンを用いる予定であるが、検索結果を、今回の目的に合致した独自の手法によって再ランキングする必要がある。この再ランキングの手法が重要となるが、現在のところ選択したWordNetの語義に付与されている例文と検索したテキストに現われる対象語の周辺の語との関連性をスコア化することによってランキングすることを検討している。評価用の100語についてランキングの精度が上がるようにスコアリングの方法を検討する。

(3) 正解の選択肢の生成

(2)の手法が確立されれば、対象語のある語義と対象語がその語義で使用されているテキストの組が得られる。正解の選択肢はこの語義と同義の語義を持つ語を選ぶことによって生成できる。正解の選択肢はWordNetに記述された同義語あるいは、語の定義から選択するが、この際に問題としての適切性を考慮する必要がある。たとえば、"tentatively"と同義の選択肢として"in a tentative manner"のように文字列が重複するものを正解の選択肢として使用するのとは"tentatively"の意味を理解していなくても選択できる可能性があり、正解の選択肢としては適切ではない。(1)の分析結果をふまえて、このような文字列の一致の他に、問題の適切性としてどのような点を考慮すべきかを明らかにし、それを考慮した上で正解の選択肢を生成する手法を検討する。

(4) 不正解の選択肢の生成

(1)でおこなった問題分析の知見、特に正解と不正解の選択肢の関連性をふまえて、不正解の選択肢を生成する手法について検討する。ひとつの可能性として、文脈中で対象語と共起している語の類義語などを不正解の選択肢として使うことを検討している。

(5) 自動作成した問題の主観的評価

(1)から(4)の研究項目の結果をふまえて語義選択問題を自動作成するプロトタイプを構築し、人手で作成された既存の問題と同じ対象語について問題を作成し、英語教師による主観的な評価をおこなう。そのフィードバックをふまえて各研究項目の内容を洗練する。

(6) 自動作成した問題の客観的評価

自動作成した問題を英語学習者に解かせ、自動作成した問題を評価する。学習者の数は100名程度、問題数はひとりあたり60問程度

を予定している。まず、対象語の集合について人手で作成された問題と自動作成した問題を用意する。それぞれの問題集合を半分に分け、対象語が重複しないように、半分ずつを入れ換える。学習者を半分に分け、それぞれに別の問題集合を割当てる。これにより、各学習者は人手作成と自動作成の問題を半数ずつ解くことになる。評価の分析は、問題の信頼性、妥当性、識別力の観点から項目応答理論モデルを基におこなう。

4. 研究成果

英語語彙問題を自動作成する以下のような手法を提案した。英語の語彙問題は(1)対象となる語、(2)対象語が使われている短いテキスト、(3)テキスト中の対象語と同じ意味を持つ語（正解の選択肢）、(4)テキスト中の対象語とは異なる意味を持つ語（ディストラクタ）の4つの要素から構成されている。解答者はテキストを読み、選択肢の中からそのテキスト中で使われている対象語の意味と同じ意味を持つ語を選択肢の中から選ぶことになる。

提案手法では、入力として対象語とその対象語の語義を与え、対象語がその語義の意味で使われているテキストをインターネット中から検索し、辞書から対象語の指定された語義と同じ意味を持つ語、および違う意味を持つ語を選択することによって語彙問題を構成する。ここで難しいのはインターネット中から指定された語義で使われている対象語を含むテキストをどうやって見つけるかという点である。テキスト中の語の語義を同定する課題は言語処理の分野では「語義のあいまい性解消」と呼ばれており、長い研究の歴史があるが、テキストによっては高精度で語義を同定することは困難だと言われている。本手法では与えられたテキスト中の語義を同定するのではなく、与えられた語義を含むテキストを探索することによって、語義あいまい性解消の困難さを回避している点が重要な貢献である。言い方を変えると語義あいまい性の簡単なテキストを選択することによって語義の同定精度を上げているともいえる。

予備実験の結果では、典型的な語義あいまい性解消の手法の精度が60%程度であるのに対して、提案手法では95%を越える精度で指定された語義とその語義の意味で対象語

が使われているテキストを取得することができた。提案手法によって自動作成した語彙問題を7名の英語教師が評価した結果、半数以上の問題が十分な質であるという判定を受けた。

さらに提案した英語の語彙問題の自動作成手法をシステムとして実装し、システムで作成した質問の評価実験を中心に研究をおこなった。対象語とその語義の1つを入力として、システムは文章中の対象語に最も近い意味を求める多肢選択問題を作成する。評価にあたっては、作成した問題が(1)英語学習者の語彙力をどの程度正しく測定できるか、(2)人間が作成した問題と比べてその質は十分かの2つの側面を考慮して評価をおこなった。

評価(1)では、自動作成した問題と人間が作成した質問をそれぞれ25問用意し、これを79名の大学生に解かせ、各学生の英語標準テスト(TOEIC, TOEFL, CASEC)の得点と比較し、両者の関係についていくつかの観点から分析をおこなった。まず、自動作成した問題と人間が作成した問題の成績には0.63と正の相関があることを確認した。次に古典的テスト理論に基づき、difficulty indexとdiscrimination indexを計算して項目分析をおこなった。さらに、Neural test理論に基づき、item category reference profileを正解、ディストラクタについて計算した。その結果、いずれの理論においても自動作成した問題が学生の英語語彙能力を測る上で遜色がないことがわかった。

評価(2)では自動作成した問題と人間が作成した問題を25問ずつをランダムに混合し、8名の英語教師に主観評価をさせた。評価項目としては、機械によって作成された問題と人間によって作成された問題をどの程度区別できるか、その判断の根拠となる問題の特徴は何か、問題が実試験に使用可能かといった観点を用意した。その結果、25問中18問についてはそのまま実試験に使用可能であるという判定を得た。

以上、提案手法によって自動作成された問題は、英語の語彙能力を測定する上で人間が作成した問題と同等のレベルであることを確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕 (計1件)

(1) Yuni Susanti, Takenobu Tokunaga, Hitoshi Nishikawa and Hiroyuki Omari, Evaluation of automatically generated English vocabulary questions, Research and Practice in Technology Enhanced Learning, Vol.12, No.11, 21 pages, Springer, DOI: 10.1186/10.1186/s41039-017-0051-y, 2017, Mar. (査読あり)

〔学会発表〕 (計3件)

(1) Arief Yudha Satria and Takenobu Tokunaga, Automatic generation of English reference question by utilising nonrestrictive relative clause, Proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017), Porto (Portugal), pp. 379-386, 2017. April. (査読あり)

(2) Yuni Susanti, Hitoshi Nishikawa, Takenobu Tokunaga and Obari Hiroyuki. Item difficulty analysis of English vocabulary questions, Proceedings of the 8th International Conference on Computer Supported Education (CSEDU 2016), Rome (Italy), pp. 267-274, 2016. April. (査読あり)

(3) Yuni Susanti, Ryu Iida and Takenobu Tokunaga, Automatic generation of English vocabulary tests, Proceedings of the 7th International Conference on Computer Supported Education (CSEDU 2015), Lisbon (Portugal), pp. 77-87, 2015. May. (査読あり)

〔図書〕 (計0件)

〔産業財産権〕

○出願状況 (計0件)

○取得状況 (計0件)

〔その他〕

6. 研究組織

(1)研究代表者

徳永 健伸 (TOKUNAGA, Takenobu)

東京工業大学・情報理工学院・教授

研究者番号：20197875

(2)研究分担者

小張 敬之 (OBARI, Hiroyuki)

青山学院大学・経済学部・教授

研究者番号：00224303