

論文 / 著書情報  
Article / Book Information

題目(和文)	
Title(English)	Automatic English Vocabulary Question Generation for Efficient Measurement of Learner Proficiency
著者(和文)	YuniSusanti
Author(English)	Yuni Susanti
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10995号, 授与年月日:2018年9月20日, 学位の種別:課程博士, 審査員:徳永 健伸,岡崎 直観,宮崎 純,村田 剛志,藤井 敦,宇佐美 慧
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10995号, Conferred date:2018/9/20, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

(博士課程)

## 論文審査の要旨及び審査員

報告番号	甲第	号	学位申請者氏名	Yuni Susanti		
		氏名	職名		氏名	職名
論文審査 審査員	主査	徳永 健伸	教授	審査員	藤井 敦	准教授
	審査員	岡崎 直観	教授		宇佐美 慧	准教授
		宮崎 純	教授			
		村田 剛志	准教授			

### 論文審査の要旨 (2000 字程度)

本論文は“Automatic English Vocabulary Question Generation for Efficient Measurement of Learner Proficiency”と題し、英文 7 章から構成されており、英語文章中の指定された単語の意味に一番近い語を 4 つの選択肢から選ぶ英語語彙問題を対象とし、問題の難易度を制御可能な問題の自動生成手法を提案すると同時に提案手法をコンピュータ適応型試験 (Computerised Adaptive Test: CAT) と組み合わせることによって、CAT の実施に従来不可欠であった事前試験による問題の難易度の推定を回避することを論じている。

第 1 章の“Introduction”では、本論文の背景と目的について述べている。言語学習において言語能力を測定することは不可欠であり、そのための試験問題の作成は専門家が人手でおこなっている。この問題作成の自動化が本論文の目的のひとつであると述べている。また、近年では、試験をコンピュータ上でおこなうことにより、受験者が同一の試験を受けるのではなく、受験者の能力に応じて、問題の難易度を動的に選択し、効率的かつ正確に能力を測定する CAT が実用化されている。しかし、CAT では事前試験によって各問題の難易度を推定する必要があり、これにはコストおよび問題の漏洩という課題が残されている。これを解決するために、問題の自動作成時に問題の難易度を制御することが本論文のもうひとつの目的であると述べている。

第 2 章“Related work”では、本論文に関連する過去の研究を(1) 問題の自動生成、(2) コンピュータ適応型試験、(3) 問題の自動生成と適応型試験の統合、(4) 問題の評価手法の 4 つの観点に分類して、それぞれの先行研究を紹介することによって、本論文の位置付けを述べている。

第 3 章“Automatic question generation”では、本論文で対象とする問題が(1) 対象語、(2) 対象語が出現するパッセージ、(3) 正解、(4) 誤選択肢の 4 つの要素からなることを説明し、それぞれの要素の作成方法を順に述べている。特にパッセージ中の対象語の語義を同定することは、正解および誤選択肢の作成に必要不可欠であるが、現在の自然言語処理の技術では文章中の語義を正確に同定することが難しく、これを解決するために文章中の語義を同定するのではなく、指定された語義の意味で対象語が使われているパッセージをインターネットから検索する手法を提案している。提案手法によって生成した 50 問と人手で作成した 50 問を混合し、英語教師による主観評価、大学生 79 名に試験を実施した結果による評価の 2 つの観点から評価し、いずれの観点からも人間が作成した問題と比べて遜色のない質の問題が自動生成できることを確認している。

第 4 章“Distractor improvement”では、第 3 章での評価結果をふまえ、さらに改善すべき要素として誤選択肢に注目し、より適切な誤選択肢を作成する手法について述べている。従来法が正解と誤選択肢の意味的な類似性のみに着目して誤選択肢を作成していたのに対し、提案手法では、これに加えて正解からできるだけ意味的な距離が大きく、パッセージ中の周辺文脈と共起しやすい語を誤選択肢として選んでいる。人間が作成した 45 問について、そのパッセージと対象語、および正解を使って従来手法と提案手法でそれぞれ誤選択肢を作成し、合計 135 問 (45 問×3 セット) を用意し、これらをプロの問題作成者による主観評価と大学生 80 名への試験実施結果の分析によって評価した結果、従来手法

よりも提案手法による誤選択肢が人手による誤選択肢に近い評価を得たと述べている。

第5章“Controlling item difficulty”では、問題の難易度に影響する要因として、(1) 対象語の難易度、(2) 正解と誤選択肢の意味的類似性、(3) 誤選択肢の語の難易度の3つを想定し、これらの要因を制御することにより、問題の難易度を制御する手法を提案している。各要因について2水準を設け、8種類の難易度を持つ問題を24問ずつ、合計192問作成し、この試験を高校生116名に実施した結果、8種類の難易度の問題に対する正答率には統計的に有意な差が認められ、提案した3要因によって問題の難易度を制御することが可能であることを確認している。

第6章“Integration of AQG and CAT”では、第5章で提案した問題の難易度の制御手法をCATと組み合わせることによって、事前試験による各問題の難易度の推定をおこなわずに受験者の習熟度の測定がどれくらい正確に可能であるかを第5章で収集した実験データを用いてシミュレーション実験で確認している。実験の結果、提案手法は事前試験による推定結果を用いた結果よりは測定誤差が大きいが、受験者全員が同一の試験を受ける一斉試験よりは測定誤差を抑えることができたと述べている。

第7章“Conclusions”では、本論文が選択式英語語彙問題の自動生成の分野に貢献した項目をまとめ、結論として要約している。特に問題の難易度を制御できる自動生成手法およびそれをコンピュータ適応型試験に組み込むことを提案したことが本論文の貢献であると述べている。また、本論文の研究成果をさらに発展させるための方向性として、対象とする問題の種類的一般化、本手法をコンピュータ適応型試験に組み込んだシステムの実証実験などを示している。

以上要するに本論文は、選択式英語語彙試験を自動生成する新しい手法を提案し、実データによる評価を通してその有効性を示している。本論文の成果は、試験問題の自動生成のための基盤技術として位置付けることができ、工業上貢献することが大きい。よって本論文は、博士(工学)の学位論文として十分価値があるものと認める。

注意：「論文審査の要旨及び審査員」は、東工大リサーチポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。