

多肢選択問題における Open Cloze 問題への適用と複数正答の可能性 Towards Open Cloze Reformulation of Multiple-Choice Vocabulary Questions: A Study on the Validity of Allowing Multiple Correct Answers

北川 みやび^{*1}, 竹中 要一^{*1}
Miyabi Kitagawa^{*1}, Yoichi Takenaka^{*1}

^{*1} 関西大学総合情報学研究科

^{*1} Graduate School of Informatics, Kansai University

Email: k189216@kansai-u.ac.jp

あらまし: 本研究では, 英検 2 級の選択肢付き語彙問題を Open Cloze 問題に転用可能かを検討した. BERT の Masked Language Model を用いて空欄語を予測した結果, 正答語が Top-10 に含まれる割合は高く, 類義語も多く出力された. これは, 文脈に応じた複数正答を許容する語彙学習支援の可能性を示唆する.

キーワード: BERT, Open Cloze, 語彙学習, 多肢選択問題

1. 背景と目的

英語学習において語彙習得は重要であり, 語彙サイズ拡大に向けた学習法が検討されている. 近年の研究では, 類義語を整理したり, 単語が用いられる文脈の中で意味を理解して記憶するなどの「深い処理」を伴う語彙学習方略が, 語彙サイズと正の相関を持つことが示されている⁽¹⁾. これらは, 意味理解を促し, 語彙の定着・拡張に有効とされる.

中高生にとって深い処理は有効とされるが, 実際の学習は反復暗記が中心である⁽¹⁾. この現状を踏まえ, 深い処理を自然に促す方策が求められている.

深い処理を学習者に促す手段の一つとして, 出題形式に着目した研究も行われている. たとえば, I always ___ up early. のように, 日本語訳を提示しない英文の空欄補充問題では, 学習者が文脈から語の意味を推測しようとする認知的処理が促されやすいことが指摘されている⁽²⁾. この形式は, 文脈から意味を理解させる構造を持ち, 深い処理を促すと考えられる.

空欄補充問題は, 語彙の知識を測定する手段として広く用いられており, 選択肢から正解を選ぶ形式が一般的である. 一方で, 空欄に学習者自身が語を記述して補う自由記述の形式 (Open Cloze 問題) も注目されている. この形式は, より多様な語彙運用力を測定できる可能性がある⁽³⁾. しかし, Open Cloze 問題には, 正解が一つに定まるよう精緻な設計が必要で, 作問者にとって大きな負担となる. そのため, 教育現場への導入は容易ではないという課題もある.

本研究では, Open Cloze 問題の作成にかかる負担を軽減する手法として, 既存の選択肢付き空欄補充問題を Open Cloze 問題として適用できる可能性を検討する. 具体的には, 選択肢を除いた状態でも適切な語が文脈から推測可能かを検証することで, Open Cloze 問題をより実践的に導入するための手がかりを得ることを目的とする.

2. 研究手法

本研究では, 選択肢付き問題を Open Cloze 問題に適用できる可能性を検討する. そのために, BERT の Masked Language Model (MLM) で, 空欄に入る語を予測した⁽⁴⁾.

使用したモデルは, transformers の bert-base-uncased である. データは, 実用英語技能検定 (英検) 2 級の空欄補充問題のうち, 単語の意味を問う 30 問を英検協会の公式ウェブサイトで開催されている過去問から選定した.

各問題に対して, 空欄部分を [MASK] トークンに置き換えたうえで, 次の 2 通りの処理を行った. 1 つ目は, 正答を含む 4 つの選択肢それぞれについて, その語が [MASK] に入る確率を, BERT の出力スコアに基づいて取得した. 4 語のうち最も高い確率の語を BERT の選択とし, 正答率を算出した. この処理は BERT が空欄補充問題に対して妥当な語を選択できることを確認し, 選択肢なしでの出力結果の信頼性を補強するために行った.

2 つ目は, 選択肢を与えずに, 問題文のみを BERT に入力し, モデルが予測した語の上位 10 件とそれぞれの確率を取得した. 出力が正答や意味的に妥当な語と一致するかを確認し, Open Cloze 問題として使用可能かどうかを検討した.

3. 結果

3.1 選択肢付き形式の出力結果

各問題において, BERT が選択肢の中から最も高い確率を出力した語が正当と一致したかを評価した. その結果, 30 問中 28 問で正答を選択した. 残りの 2 問は正答が複数のトークンに分割されたために確率の算出ができなかった. この 2 問を除く 28 問ではすべて正答を選択しており, 実質的な正答率は 100% であった.

3.2 選択肢なし形式における出力結果

選択肢なしで [MASK] を含む文のみを BERT に入力した結果, 正答語が Top-1 に出力されたのは 5 問 (16.7%), Top-5 では 14 問 (46.7%), Top-10 で

は17問(56.7%)であった。

一方で、Top-10に正答語が含まれなかった13問の中にも、文脈上自然な語が出力された例もあった。

表1には、正答語がTop-1に出力された例と、正答語が出ずとも妥当な語が上位に出力された例を含む。

表1 BERTの出力例

問題文(抜粋)	正答語	出力上位3語(確率)
...to [MASK] a new show...	create	create (24.8%) make (18.9%) produce (17.2%)
...to [MASK] a cold.	cure	treat (54.1%) cure (26.8%) prevent (3.4%)
...could not [MASK] many...	grasp	understand (43.0%) read (12.7%) explain (7.0%)

4. 考察

本研究では、選択肢を除いた選択肢付き問題でも、BERTが文脈に応じて妥当な語を出力できるかを検討した。その結果、正答語がTop-10に含まれた問題は約6割にのぼり、特にTop-1に正答語が出力された例も見られた。一方で、正答語が出力されなかった問題においても、“grasp”に対して“understand”のように、意味の近い語が高確率で出力される例が多く確認された(表1)。これらの出力傾向は、語彙学習の文脈において、選択肢付き問題をOpen Cloze問題に適用する可能性を示唆する。

正答語でなくても意味的に関連した語が上位に出る傾向は、語彙学習支援において重要な示唆となる。語の意味を文脈に基づいて判断し、複数の関連語を想起・整理する過程は、深い処理を伴う学習につながる⁽¹⁾とされており、BERTの出力がこのような処理を自然に促す構造になっていると考えられる。たとえば、“grasp”が思い出せない場面でも、“understand”や“comprehend”などの語が提示されることで、意味のまとめりとしての記憶や関連語の体系化が可能になる。この仕組みは、語彙運用の柔軟性を高める支援となる。

さらに本研究では、正答語が出力されなかった場合においても、文脈上妥当とみなせる語が上位に出力されるという傾向が一貫して観察された。これは、語彙学習において「複数の正解」を前提とする柔軟な理解や受容が有効であることを示している。語彙は本質的に多義性や使用文脈の幅を持ち、単一の語に正解を限定する形式は、かえって学習者の語彙運用の柔軟性を損なう可能性もある。

従来の選択肢付き問題では、採点や指導上の都合から「唯一の正解」を設ける必要があったが、学習支援を目的とする場合には、文脈に適合する複数の

語を受け入れる設計の方が、語彙間の意味的ネットワークを理解させ、深い処理を引き出す教材として効果的である。

5. まとめと展望

本研究では、既存の選択肢付き空欄補充問題を、選択肢なしのOpen Cloze問題に適用できる可能性を検討することを目的とし、BERTのMLMを用いて、文脈に応じた語の出力傾向を検討した。その結果、正答語がTop-10に含まれる割合は約6割にのぼり、正答語が出力されなかった場合でも、文脈上妥当とみなせる語が高確率で出力される例が多く確認された。

これらの結果から、BERTの出力は、語彙学習において文脈に応じた複数の語の受容を支援する仕組みとして活用できる可能性が示唆された。従来のように「1つの正解」のみに限定するのではなく、意味的に近い語を含めて柔軟に受容できる設計は、語彙の類義語などを意識させ、深い処理を促す語彙学習の実現に寄与することが期待される。

一方で、本研究はあくまでBERTの出力傾向を分析したものであり、実際に学習者がこの形式の支援によってどのような効果を得るかについては検討していない。また、「意味的に妥当である語」の判断は、主に直感的評価に依拠しており、語彙間の意味距離や語用論的な制約を定量的に評価する枠組みは未整備である。

今後は、学習者の反応や語彙習得の定着度を含めた実証的研究を通じて、複数解を許容する語彙学習支援の有効性を検証することが求められる。さらに、ターゲット語の定着を妨げることなく、類義語との違いや使い分けに焦点を当てた提示方法の設計、ならびに語彙レベルや学習段階に応じた動的な出力制御の仕組みについても検討する必要がある。

参考文献

- (1) 内田奈緒: “中高の英語学習における語彙学習方略 — 方略使用・有効性と規定要因に関する発達の差異の検討 —”, 教育心理学研究, 第69巻, 第4号, pp.366-381 (2021)
- (2) 境奈津希: “どのようなテスト形式が高校生の語彙学習を促すのか”, 中部地区英語教育学会紀要, 第49号, pp.133-138 (2020)
- (3) 松森匠哉ら: “マスク言語モデルを利用した Open Cloze 問題の自動生成”, 人工知能学会全国大会論文集, 第36回, セッション3N4-GS-10-04 (2022)
- (4) Devlin, J., et al., : “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”, Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol.1, pp.4171-4186, Minneapolis (2019)