

博士論文審査及び最終試験の結果

審査委員（主査） 佐野 洋 印

学位申請者 川本渚凡

論文名 Determining Factors of Good Example Selection for EFL Teaching and Learning: Toward Methodologies and Evaluation Criteria.

【審査結果】

佐野洋を主査とし、主任指導教員の投野由紀夫、および副査として根岸雅史、吉富朝子、山田茂（外部委員、早稲田大学）から成る審査委員会は、2024年6月22日に上記論文の審査ならびに口述による最終試験を行った。その結果、審査委員会は全員一致で、申請者に対し博士（学術）の学位を授与するのが適当であるとの結論に達した。

【論文の概要】

本研究の目的は、英語教授・学習に有効な例文の選定・評価基準の理論的枠組を構築し、その内部の言語的な特性と外部の諸条件との関連性を実証的に調査することを目的としている。例文は、教科書、文法書、単語帳など、さまざまな言語学習教材で一般的に提供されるが、その例文を評価するための確立された基準は存在していない。辞書学、自然言語処理（NLP）、コンピュータ支援言語学習（CALL）、データ駆動型学習（DDL）などの分野では、これまでも「良い」例文の基準に関する多くの研究が行われてきたが、その多くが、例文難易度決定に寄与する言語的特徴にのみに焦点を当てている。一方で、例文の質を決定する要因は、この言語的特徴以外にも複数存在する。しかし、これまでに例文の質を決定する種々の要因に総合的に焦点をあてた実証研究は、ほとんど存在しない。

本論文ではまず全体の序章の後、第2章で辞書学、NLP、CALL、DDL分野の先行研究のレビューを通し、例文選定の条件や例文使用の効果に寄与する変数整理を行った。それらの情報をもとに第3章では「例文の使用と評価のモデル」と称して暫定的な枠組みを提案した。このモデル内では、例文選定・使用を評価する要因として、例文が提供する言語的な情報（[INFORMATION]）、その例文を参照するユーザー（[USER]）、例文が提供される媒体（[MEDIUM]）、例文が提供される目的（[PURPOSE]）の関係を相互に位置付けている。これらの要因のほか、例文評価の基準（CRITERIA）は、USER, MEDIUM, PURPOSEの相互作用で決定され、それが用例の有用度を定めるものであるとする。例文評価の基準のうち、言語的な特徴を基準言語特性（[Criterial Linguistic Properties]）と呼び、これは例文が与え

る情報の種類を定義する INFORMATION とは区別している。INFORMATION は例文がその用途として提示するさまざまな情報であり、基準言語特性は例文の質決定に寄与する言語的特徴としている。これらの枠組を用いて、本論文では以下の 3 つの研究設問を設定した：

RQ1：例文の質・難易度を評価する基準言語特性の性質とは何か

RQ2：例文の基準言語特性はその他の関連要因（目的・ユーザー・媒体）に基づき、いかに重み付けされるか

RQ3：例文は特定の条件下で実際のユーザーにどのように使用されるか

これらの研究設問に応えるべく、本論文では 4 つの実証研究を行い、それらの知見をもとに第 3 章の「例文の使用と評価のモデル」の改訂を第 8 章で試みた。まず 3 つの実証研究（第 4 章から第 6 章）により、上記の諸要因間の相互関係を調査した。第 4 章では、基準言語特性と例文提供目的の関係([**Criteria** Linguistic Properties] × [PURPOSE])、第 5 章では、基準言語特性と例文が想定している難易度レベルの関係 ([**Criteria** Linguistic Properties] × [USER])、第 6 章では、基準言語特性に基づき設定された例文が実際の学習者にとってどのように認識されているか ([**Criteria** Linguistic Properties] × [USER] ↔ [REAL USER])に着目した。これらの結果を踏まえ、第 7 章では、前章までに明らかとなった例文の質に寄与する言語的特徴に着目し、統合的に例文使用の観察調査を実施した。

第 4 章では、例文を提示する媒体としての辞書に着目し、目的別辞書の例文を標本抽出し、それらに特徴量として予想される言語的特徴（文長、句例 vs 文例、語彙レベル等）を例文単位で注釈付けし、例文提供目的（発信用辞書 vs 受信用辞書）を区別する基準言語特性を、条件付推論ツリーモデル（**conditional inference tree**）を用いて探索した。その結果、特に 2 言語（**bilingual**）辞典において、例文の形式的特性（句例 vs 文例）が、受信用例文と発信用例文のタイプを弁別するのに最も寄与することが判明した。一方、完全文の用例を提供する傾向の強い 1 言語（**monolingual**）辞書では、用例の長さ（例文中に含まれる単語数）が受信用と発信用の例文タイプの決定に寄与していた。これらの 2 つの要因に加えて、例文の平均語彙レベル、文内に含まれる B1 レベルおよび B2 レベルの単語の割合等の語彙的特徴も部分的に決定に寄与することがわかった。

第 5 章では、例文の難易度に寄与する言語的特徴を探索するため、CEFR のインベントリ（**Core Inventory for General English, Global Scale of English, English Grammar Profile**）で提示されているレベル別例文を調査対象とし、これらの例文の基準言語特性によって、例文難易度がどのように規定されるか、順序ロジスティック回帰分析を用いて調査した。結果は、インベントリの種類に関わらず一貫して例文の難易度に寄与する言語的特徴は限定的であった。すなわち、用例内の A1 レベルの単語数、統語構造の深さ、用例に含まれ

る名詞の数、用例の平均文法レベル、そして用例内に含まれる A2.2 レベルの文法構造の数が例文難易度に寄与する変数として特定された。一部の特徴量を除き、これらの基準言語特性は順序ロジスティック回帰モデルで CEFR レベルの上昇に対し正の係数となり、レベル上昇の特徴量として有効であることが示された。

第 6 章では、第 5 章で取り上げたインベントリーに収容された例文を実際のユーザーが読んだ際に感じる難易度の感覚を実験的に調査した。その意図は、専門家によって作成された CEFR インベントリーは難易度に寄与する特徴について重要な知見を提供している一方で、学習者がこれらの例文の難易度をどのように受容するかについては、必ずしもインベントリーで想定された難易度と一致するとは限らないからである。そのため、Core Inventory for General English と Global Scale of English を対象に、学習者が感じる難易度レベルを調査した。ここでは学習者が感じる例文難易度を当該例文の和訳困難度と定義し、この学習者が感じる難易度レベルとインベントリーで規定・付与されている例文の難易度レベルに相関があるかどうか、ラッシュ分析で調査した。その結果、学習者が感じる難易度は一般にインベントリーの例文の注釈付けされた難易度レベルと一致する傾向が確認できた。ラッシュ分析によって決定されたアイテムの難易度順と相関が見られた言語的特徴は、単語数 (A2 レベル、B2 レベル)、固有名詞の有無、および構文特性の深さや例文の長さなどであった。さらに、質的分析は、注釈付けされた言語的特徴以外にも、学習者の例文の受容難易度レベルに寄与する言語的特徴が他にもあることを明らかにした。それらには動詞の時制や多義語などが含まれていた。

第 7 章では、4 章から 6 章までで明らかとなった例文の質に寄与する言語的特徴に着目し、統合的に例文使用の観察調査を行った。ここでは、基準言語特性として 6 つの変数 (句例 vs 文例、用例の長さ、用例内での見出し語の位置、用例内の難語の有無、用例の文法構造)、例文提供インターフェイスに関わる特徴として 5 つの変数 (用例数、提示順、文法コードの有無、日本語訳の有無、提示方法 (KWIC vs. 文提示)) を取り上げ、これらの変数が日本人英語学習者の辞書使用に与える影響を統合的に観察した。観察タスクとして、与えられた辞書情報をもとに、英文完成タスク、エラー修正タスク、和文英訳タスクの 3 種類の産出タスクを行った。この際、学習者の認知プロセスを確認するため、誘出方法として思考発話法を用いた。また、提供する辞書情報は、上記の基準言語特性と例文提供インターフェイスの変数の影響が確認できるように例文セットを作成し、セット別に学習者の遂行状況を観察した。観察の結果、基準言語特性・インターフェイスに関わる変数間で、例文使用の難易度に対して相互に複雑な影響が見られた。

第 8 章は、この論文のために行われた 4-7 章の個別研究の結果を考察しつつ、3 章で提示した「例文の使用と評価のモデル」を更新し、例文とその評価基準の役割をより確固たる理論的枠組みの中に位置づけることを目指した。その中で、本研究が方法論的には混合研究法 (mixed method) の手法を用いて辞書例文の目的・条件・評価基準の体系化を試みた

点、理論的には辞書学・自然言語処理などの諸分野における例文研究に一定の理論的な枠組を提案した点、教育的には特に目的別・対象ユーザー別の例文のあり方を明確にし、それに基づくより有効な例文作成・選定の実際的な方法の提案に道を拓いた点等、本論文の外国語教育学・辞書学分野での貢献を述べ、特に辞書編纂の分野で将来的にレベル制御された例文を提示するインターフェイスの改善などの可能性を論じた。

【審査の概要及び評価】

審査では、川本氏による博士論文の概要の発表の後、各審査委員から本論文の評価できる点として、以下のようなコメントがあった。

1. 「例文・用例」の質の評価の枠組を理論的に整理し、関連する内的・外的要因の複雑な相互関係を量的・質的な実証データをもとにその性質や評価基準に迫ろう、という意味で外国語教育分野での初の本格的な「用例」研究だといえる。
2. 特に例文の言語特性の分析では、計 5,330 例に対して文単位の形式・語彙・文法プロファイル（構文解析による構文木の深さ計測など）の詳細なアノテーションを行っており、このデータセットは中間報告を欧州において発表した際にも高い関心を持たれ、CEFR レベル例文判定研究などにも応用できる資源で価値が高い。
3. 複数の用例リストをコーパスとして用いて、これら実データから妥当性を計算する各種の数理モデルを利用して量的に担保された実証研究を行っていること。例えば、上述のデータタグ付けの労力だけでも評価に値するが、回帰分析を用いて CEFR レベルの決定に寄与する言語的特徴を子細に明らかにした点は高評価に値する。
4. 定量分析に加えて、利用者である学習者視点の測度をモデルに反映するため、思考発話法（think aloud）を用いた定性分析を行っていること。被検者規模も大きく妥当性を有し、用例の学習者適合に関わる問題点を数多く指摘しており、学習者の特性と例文の特性との関係を十分に調査することが不可欠であることを実証的に示した。
5. 例文使用のさいの意図された目的（理解 vs 産出）を変数として、言語特徴などの要因との関係を明らかにしたことは従来研究では見落とされがちだった観点で、注目に値する。

このように、川本氏の博士論文が高く評価すべきものであると確認された上で、質疑応答では以下のような指摘およびそれに対する回答があった。

1. 本研究では一貫して因子を「探索的」に分析していたが、かたや「例文とはこうであるべきだ」という知見に基づいた仮説検証的なアプローチの分析もありえたのではないか？

これに対して川本氏は先行研究の整理の段階で、「べき論」としての例文の資質・条件はある程度整理したつもりであったが、モデルそのものにもう少し顕在化することもできたかもしれない、と回答した。

2. 単語難易度は **CEFR-J Wordlist** に基づいているが、用例内の単語の異なる語義についてはレベル表示ができない。また成句を考慮していない、など例文の意味解釈の面ではより緻密な分析が必要だったのではないか。

これに対して川本氏は現状の単語プロフィール研究がまだ語義や成句に踏み込んだレベル設定をしていないものがほとんどだったので分析は見出し語レベルとしたが、今後は用法や意味も考えていきたい、と回答した。

3. 用例の評価モデルとしては有効だと思うが、辞書検索プロセスにおける用例の位置づけやその各段階での使い方や役割なども明らかにして欲しかった。

これに対して川本氏は、今回は辞書以外に **CEFR** インベントリーなどでの用例提示も対象としたため、辞書検索プロセスの具体まで踏み込めなかった、と回答した。

4. 作る側 (**material developer**) が最後のモデルに突然現れる印象だが、例文を作る辞書開発者側の編集意図や例文に対する考え方などはもう少し掘り下げ、辞書制作過程においてより明確に位置付けてもよかったのではないか。

これに対して川本氏は、一連の実証研究で例文の資質・条件を明確にしつつ、改訂版モデルで開発側とユーザーの目的・意図の乖離を示したつもりであったが、論文の最初のモデル提示の際に作る側の位置づけをより明確に論じた方がよかったかもしれない、と回答した。

5. **Inventory** の例文は選定の目的および例文選定ソース（学習者データか教科書か、など）によっても違いがあると思うので、一律に **CEFR** レベルを代表した例文といえるかは注意が必要だと感じた。

これに対して川本氏は、たしかにソースの性質、選定基準などが **Inventory** によって異なることは考えられるので今後はそのような違いによる例文の代表性なども勘案するべきだと思う、と回答した。

6. 7章で思考発話法を用いて観察調査を行ったが、分析結果がまとまっては提示されておらず、複合的な変数の相互関係の記述を主に取り上げていて、質的な分析が多少物足りなさを感じた。

これに対して川本氏は、7章の観察調査のデータは多くの変数を組み込んだため、その変数とタスク達成度の関係を見るのに時間を要し、発話思考法の分析は極めて限定的だった。今後の研究ではさらに詳しい分析を試みたい、と回答した。

その他、語の定義と用例の関係、文例に対する句例の位置づけ、書き言葉と話し言葉における例文の違い、AI が普及して電子辞書さえ引かなくなった世代への提言は何か、など

の質問も出た。川本氏は自身の研究の焦点を改めて確認しながら、これらの諸点に関して
も用例選定の評価モデルの意義と限界、今後の課題と展望を述べつつ真摯に応答した。

最終試験後の審査委員会での審議においては、委員からの指摘の多くが、川本氏の論文
から導かれる今後の研究の発展の可能性について指摘したもので、本調査がもつ学術的・
実践的意義を高く評価した上での期待の表れとみなすべきものであったことが確認され
た。

以上の論文評価および最終試験での質疑応答の内容から、本論文は英語教育学・英語辞
書学の双方に大いに貢献する秀逸な理論的・実証的研究であり、学位申請者が優れた研究
者としての資質を十分に有していると判断された。よって審査委員会は、全員一致で、学
位申請者が博士（学術）の学位を授与するにふさわしいとの結論に達した。