

論文審査の結果の要旨および担当者

報告番号	※	第	号
------	---	---	---

氏 名 阿部 大輔

論 文 題 目

A Data-driven Approach to the Development of Constructions in Japanese EFL Learners

(日本人英語学習者のコンストラクションの発達へのデータ駆動型アプローチ)

論文審査担当者

主 査	名古屋大学	教授	杉浦正利
委員	名古屋大学	教授	山下淳子
委員	名古屋大学	准教授	大島義和

論文審査の結果の要旨

1 本論文の概要と構成

本論文は、日本人英語学習者が英語の言語形式と意味とのマッピングであるコンストラクションをどのように習得していくかという問題に、包括的なデータ分析から解明を試みるという研究である。

本論文は、6章から構成されており、添付資料として、分析対象として使用した英文エッセイの題目一覧と、統語解析プログラムで使用されている分析タグの一覧と、構文部分木の抽出に使用したスクリプトが添えられている。

第1章(序論)では、本研究の目的と論文の構成が述べられている。主たる研究目的は、英語を外国語として学ぶ日本語母語話者が、いかにコンストラクションを習得するのかという問題の解明に取り組むことである。加えて、方法論的目的として、大規模なコーパスデータから文法構造の構成素を抽出するにあたり、これまでの先行研究では使われてきたことのない自然言語処理ツールを活用した分析手法を導入し、斬新な方法論の確立に向けた試みを行うことが述べられている。

第2章で、まず本研究の出発点となる Biber ら(2011)による第二言語習得における発達段階説とその後の検証研究を説明し、本研究の基本概念となる用法基盤理論に基づく言語習得、構文文法、連語表現と語彙バンドル、文脈自由文法と木置換文法に関して、先行研究のレビューを行っている。そして本研究でいかにそれらの概念を用いてアプローチするかを述べた後、本研究の研究課題をまとめている。Biber ら(2011)は、従来の研究が従属節など節を単位として文の複雑性の発達研究を行ってきたことに対し、アカデミックライティングでは節ではなく名詞句の複雑性が重要であることを指摘し、話し言葉と書き言葉のコーパスデータの分析から、コンストラクションの複雑さの発達段階を提案している。それに対し、母語話者の話し言葉と書き言葉の関係と第二言語における発達段階は同じではないという Yang (2013)の指摘や、Biber ら(2011)の提案した発達段階を実際に学習者データで検証した Parkinson & Musgrave (2014)、Ansarifar ら(2018)、Kreyer & Schaub (2018) 等の研究があるが、これらの研究結果は必ずしも一致していない。本研究では、Biber ら(2011)も含めてこうした先行研究が、特定の文法構造の表現のみをコーパス中で調査し比較しているため分析結果に偏りが生じることを指摘し、コーパスデータ全体で使用されているすべての文法構造の構成素を分析対象としたボトムアップ的アプローチの必要性を主張し、その分析に取り組むことが本研究の課題であると述べている。

第3章では方法論を述べている。本研究で使用するデータ(ETS Corpus)、構文解析器(Stanford Parser)、不連続データ指向構文解析器(disco-DOP)、構文木正規表現検索(Tregex)について説明をした後、先行研究と同様のトップダウン的アプローチと、本研究で新たに試みるボトムアップ的アプローチについて、分析手順を説明している。使用するデータは、母語による違いが影響しないように ETS Corpus のうち日本語母語英語学習者が書いたエッセイに限った。Low, Mid, High の3段階の熟達度と8種類のエッセイ題目の組み合わせからなる24グループのデータセットを作成し、執筆量が影響しないように、データセットごとに語数をそろえ、熟達度ごとに約2万8千語のデータセットを作り、Stanford Parser により構文解析を行った。そして、今回分析対象とするデータに対し、先行研究と同様にトップダウン的アプローチとして、Biber ら(2011)の提案した発達段階に対応する、名詞句内の修飾関係に該当する文法構造表現を、Tregex を使って検索し頻度を調べた。次に、不連続

論文審査の結果の要旨

データ指向構文解析器 (disco-DOP) を使い、構文解析データ中で使用されている構文の構成素としての部分木のすべてを対象として、使用されている部分木の種類と頻度を調べた。

第4章では結果を述べている。まず、トップダウン的アプローチによる分析結果をまとめた後、ボトムアップ的アプローチによる分析結果を詳しくまとめている。トップダウン的アプローチでは、先行研究に基づき、予め定められた名詞修飾の種類と頻度を3段階の熟達度の間で比較している。次に、ボトムアップ的アプローチでは、部分木の包括的な出現頻度について、3段階の熟達度別に **Type** と **Token** をまとめた後、文・名詞句・動詞句という三種類の種類別に、ノード数・末端ノード数・深さ・語ノード数・非語ノード数という5つの観点から部分木の複雑さを分析した。その結果、Biber ら(2011)が主張している通り、習熟度の上昇とともに名詞句の複雑性が上がることが確認された。最後に、文・名詞句・動詞句ごとに、具体的なコンストラクションの出現パターンを観察した結果、Biber ら(2011)を含む先行研究では取り上げられてこなかったが、本研究で、コピュラ構文や名詞句内の限定詞が重要な文法構造の構成素であることが発見された。

第5章では、分析結果に基づき議論している。Biber ら(2011)およびそれに続く従来のトップダウン的アプローチによる名詞句の複雑性に関しては、前置詞句による後置修飾が熟達度が上がるにつれ増加するという点は、一連の先行研究と同じ結果が得られたが、それ以外の修飾関係は、差がみられなかったり、熟達度との分布関係がずれていたりして、統一的な結論を導き出すことは困難であることが導き出されている。その後、ボトムアップ的アプローチによる分析結果について、文全体・名詞句・動詞句の三種類の単位での部分木の特徴を、そのパターンと具体的な表現の両面からの考察し、Biber ら(2011)の提案した発達段階とは違う結果が得られるとともに、Biber ら(2011)では含まれていなかったが本研究で発見されたコピュラ構文や名詞句内の限定詞に基づき、Biber ら(2011)の提案した発達段階に対する修正を提案している。

第6章では、本研究のまとめと、限界、そして、今後の研究への提言が述べられている。すなわち、先行研究と同様のトップダウン的アプローチでは、一貫した結果は得られないこと、そして、ボトムアップ的アプローチでは、一部、名詞句の複雑性の発達という点では、Biber ら(2011)の主張通りの結果が得られたものの、本研究による網羅的な文法構造の部分木の調査により、従来、見落とされていた二つの構文パターンが発見され、それを考慮に入れた、日本語母語英語学習者のデータに基づいた新しい発達段階が提案された。

2 本論文の評価

本論文は、学位論文として以下の点が評価される。

(1) 伝統的に節を単位とした文法構造の複雑さの発達観に対し、名詞句の複雑さの重要性を指摘しそれに基づく発達段階を提案した Biber ら(2011)の研究は画期的であったが、研究のもととなるデータが母語話者の話し言葉と書き言葉の対比によるものであった点と、取り上げた文法構造が網羅的ではなく恣意的に選ばれていたという問題点があった。Biber ら(2011)の提案した発達段階を学習者デ

論文審査の結果の要旨

ータにより検証を試みた後続研究も、最初に Biber ら(2011)が取り上げた文法構造に基づき検証を行ってきた。これに対し、本研究は、日本語母語英語学習者の 3 段階の熟達度のデータを用い、Biber ら(2011)およびその後続研究が取り上げた文法構造について検証するだけでなく、対象データをすべて構文解析しそこに含まれるすべての構文部分木を対象に頻度を出すという方法を使い、恣意的ではなく網羅的にコンストラクションの発達を観察したという斬新な研究であるという点が高く評価できる。

(2) 本研究の分析により、Biber ら(2011)や、それと同様にコーパスデータに対する演繹的な検索による検証を行ってきた後続の研究では発見しえなかった、コンピュータ構文や名詞句内に限定詞を含む構造が、発達段階において重要な役割を果たすことを明らかにした点が二つ目の重要な貢献であると評価できる。

(3) 新たに発見された発達段階に関する重要な文法構造を使い、Biber ら(2011)の提案した発達段階に対する具体的な修正案を提案している点が、三つ目の重要な貢献といえる。

(4) Biber ら(2011)およびその後続研究ではコーパスデータを分析しているものの、検索作業を手動で行うという分析方法であった。しかし、本研究では、それを自動化して分析を行っている。こうした分析作業の自動化の確立により、今後、ほかのデータに対しても同様の処理を速やかに行え、その検証もできるようになる。こうした新たな分析方法の確立に向けた貢献という点でも本研究を高く評価できる。

ただし、本論文は、以下の点において今後改善すべき点があることが指摘される。本研究による分析で、新たに発見された二つの文法構造を加えて従来の発達段階を修正する案を提案しているものの、その発達段階の修正案は、今回の 3 段階の習熟度別データに基づく発達順序と従来提案されてきた発達順序とを比べて、相対的に文法構造の発達順序の整合性を考えて提案されたものである。しかし、実際に各文法構造に関する言語指標（特徴量）が具体的に何らかのスコアとして特定されているわけではないので、ある言語特徴量を持つデータが具体的にどのステージかということ特定することができない。また、今回の研究では、日本語母語英語学習者のデータを用いているので、今回の分析結果は、日本語を母語とした英語学習者の場合という限定がつく。ほかの母語の英語学習者でも同様の結果が得られるかは検証されていない。また、コーパスデータの語数統制が必要であったため、データ規模が十分大きいとは必ずしも言えず、そのために、出現頻度が低く十分に検証できなかった文法構造があったことも課題として残る。しかし、これらの点は、本研究を今後さらに発展させるための課題であり、本研究の博士論文としての価値を損なうものではない。

論文審査の結果の要旨

以上の評価により、本論文は、博士（学術）の学位に値するものであると判断する。

論文審査の結果の要旨

論文審査の結果の要旨