

令和 3 年 6 月 16 日現在

機関番号：33929

研究種目：基盤研究(B) (一般)

研究期間：2017～2019

課題番号：17H02367

研究課題名(和文)日本人英語学習者の心的語彙ネットワーク拡張・深化過程の可視化とポートフォリオ開発

研究課題名(英文) Visualization and portfolio development of the expansion and deepening process of Japanese learners of English

研究代表者

青谷 法子 (Aotani, Noriko)

東海学園大学・教育学部・教授

研究者番号：00278409

交付決定額(研究期間全体)：(直接経費) 10,160,000円

研究成果の概要(和文)：本研究では、「非対称フォン・ミーゼス尺度法」(AMISESCAL)を用いて多読活動前後の語彙ネットワーク構造の変化の可視化を試み、既習語における語彙ネットワーク構造の変化を促すためには、深い意味的処理が必要であるという知見を得た。また、またアソシエーション分析およびネットワーク分析を組み合わせた独自の英語語彙ネットワーク構造可視化プログラム(LexNetViz)を開発し、ES(A)Pの観点からの分野固有語彙ネットワークの拡張・深化過程の可視化を試みた。このようなデータを蓄積することで学習者の語彙知識発達についてのポートフォリオを構築し、フィードバックすることが可能であると示唆された。

研究成果の学術的意義や社会的意義

本研究では、「非対称フォン・ミーゼス尺度法」(AMISESCAL)を語彙習得研究に応用、また独自の英語語彙ネットワーク構造可視化プログラム(LexNetViz)を開発することで、語彙知識のネットワーク構造に焦点化してその発達過程を記述することを目指した。心的辞書の構造的性質の変化を明らかにすることで、従来からの語彙知識の量的側面に焦点をあてた研究を補完する意義があったと考える。ネットワーク構造の可視化によって語彙知識に関するより豊かな診断的情報をポートフォリオとして学習者に提供することが可能となり、こうした質的フィードバックの教育的効果は大きいと考えられる。

研究成果の概要(英文)：In this research project, we focused on explicating and visualizing the development of learners' L2 mental lexicon by investigating how words were networked before and after undertaking different kinds of reading task. The data was analyzed by using AMISESCAL (Asymmetric von Mises Scaling), a statistical model that visualizes asymmetric relations among elements on a two-dimensional map, and the results indicated that deeper semantic processing plays an important role for the reorganization of the lexical network. We also developed a computer program; LexNetViz, which enables the visualization of the English lexical network structure to investigate the expansion and deepening process of ES(A)P lexicon. It was suggested that the data accumulated with the methods we used in this project could be a useful resource to give learners feedback for their portfolio that can give them a concrete image of their lexical network development.

研究分野：心理言語学

キーワード：心的語彙ネットワーク 非対称フォン・ミーゼス尺度法 アソシエーション分析 日本人英語学習者 多読 ESP

1. 研究開始当初の背景

我々の心的辞書が“giant multidimensional cobweb”（巨大で多次元に広がるクモの巣）のようなネットワーク構造を持つとする概念モデルは広く受け入れられている（Aitchison, 2003 など）。この心的辞書概念モデルは、我々の脳が、実際に、神経のネットワーク構造から成り立っていることから発想されたものであり、個々の語彙項目（ノード）が脳の神経細胞を、それをつなぐリンクがシナプスを比喩的に写像している。語彙数の増加はノードが増加すること、それぞれの語についての知識を増やすことは新しいリンクを張ることに例えられ、このモデルは我々の脳の中で実際にそのような現象が起こっているかのようなリアリティを与えている。

実際、そうしたネットワークメタファーを、意味的に関連する語群の提示や学習定着に援用する取り組みはこれまででもなされている。しかし、多くの場合、この関連性は理論的に導かれたものであり、目標言語使用者が実際にそれらの語をどのように関連づけているか、といった実証データに基づくものは少ない。一方、語彙知識の構造的側面に対し実証的にアプローチする研究の多くは、対象とする語彙が限定されているか、語彙項目間に対称的な（両方向に等しい）関係を前提としている。しかし、学習者の語彙習得を追跡するにあたって、語彙項目間の関係性は対称的とは限らない（例えば連想関係において apple を刺激語として red を連想する比率と red を刺激語として apple を連想する比率は同じではない）だけでなく、学習者が到達すべき語彙サイズを考慮に入れると、数千語単位での語彙項目を最終的には分析の対象とする必要がある。

こうした問題意識を元に、申請者らは非対称データを分析し視覚化するための統計モデルである「非対称フォン・ミーゼス尺度法」(AMISESCAL)を用いて学習者から得られた語彙データの分析を行った。AMISESCAL は、2011 年に荘島によって開発・提唱された新しい統計モデルであり、その英語教育学の分野における応用は世界的にも先駆的な試みである。本来、非対称的である語彙項目間の関係性を、従来のように対称的なデータであるかのごとく扱うのではなく、非対称性を持つ情報量をありのまま活用できるという点で従来の分析手法よりも正確な分析とネットワーク構造の可視化が可能となった。また、グラフ可視化プラットフォームである Gephi (Bastian, Heymann, & Jacomy, 2009) を用いた語彙ネットワーク構造の可視化にも取り組んだ。さらに、アソシエーション分析を援用することで、個々の学習者から得られたのべ 1,000 語・4,500 ペア以上の語連想データからその学習者にとって十分に確立されていると考えられる共起ルールを抽出し、ネットワーク図として提示した。

これらの研究成果は、コンピュータの処理速度の向上、新しい分析手法の開発、グラフ理論やテキストマイニングの手法の応用によって可能になったものである。こうした手法を用いることで、個々の学習者がある一時点に習得している語彙知識の構造的性質やネットワーク構造を把握し、共時的・横断的な比較分析が可能となった一方で、学習者の言語経験を基盤とした語彙知識の発達過程の解明には、通時的・継続的な研究が必要となっている。申請者らはこれまでの研究成果を踏まえて十分な研究遂行の見通しが得られたため、この課題に取り組むこととした。

2. 研究の目的

本研究では、日本人英語学習者の心的語彙ネットワークの構造が拡張・深化する過程を実証的データに基づいて可視化し、その促進要因の特定とその機序を次のような手順で明らかにすることを目的とした。

- (1) 多読活動を通して語彙ネットワークが拡張・深化する過程を可視化し、その促進要因を特定し、その機序を明らかにすること。
- (2) ES(A)P の観点から特定分野に特徴的な語彙ネットワークをコーパスに基づいて目標ネットワークとして設定し、その分野の学習者がその分野固有の語彙ネットワークを構築する過程を明らかにすること。
- (3) 語彙ネットワーク構築過程を可視化するためのプログラムを開発し、学習者・指導者にポートフォリオとして提示するためのシステムを構築すること。

3. 研究の方法

学習者から得られた語と語の関係性に関するデータを AMISESCAL を用いることで語彙ネットワークとして可視化、分析を行った。また、アソシエーション分析とグラフ可視化プラットフォームである Gephi を組み合わせた独自の英語語彙ネットワーク構造可視化プログラム (LexNetViz) を開発し、ES(A)P の観点からの分野固有語彙ネットワークの拡張・深化過程の可視化を行った。これらの研究から得られた診断的情報をポートフォリオとして個々の学習者に提供する手法の確立を試みた。

4. 研究成果

- (1) 語彙ネットワーク構造の発達過程を解明するために、日本人大学生を対象に 8 つの語を提示し、次のような手順で研究を行った。8 つの語のネットワーク関係を自由に描画。8 つの語それぞれの間の関係の強さを 6 段階で評価し、その結果を AMISESCAL を用いて分析し、そこから導き出された語彙ネットワーク構造をプロット図として可視化。再び 8 つの語の関係を自由に描画、得られたネットワーク構造との比較を行った。参加者を TOEIC の得点によって英語運用能力上位群と下位群に分けて分析を行った結果、上位群において と との類似性が高いことが示された。さらにすべての参加者において は よりも複雑なネットワーク

構造を示した。本研究により、上位群ほど、1語1語の間の関係性を意識しながらネットワーク構造をとらえていること、また、

この知見を基に、多読活動にどのようなタスクを組み合わせることが学習者の心的語彙ネットワーク構造の拡張・深化につながるかという実証的データを得ることを目指し、多読前後の語彙ネットワーク構造の変化の可視化を試みた。同じ多読課題に対し異なった処理水準(level of processing)を持つタスクを与えた2群(日本語への部分訳群(RM+TL Group)、TOEIC形式の選択式問題群(RM+MCQ Group))の日本人学習者が、実験の前後で語彙間の関係をどのように捉えるのかについて実験を行った。実験の前後において、多読教材に出現するキーワードを用い、それぞれの間の関係の強さを視覚的アナログスケールで評価させ、その結果をAMISESCALを用いて分析し、導き出された語彙ネットワーク構造をプロット図として可視化した。与えられたタスクをHulstijn & Laufer (2001)によって提唱されたInvolvement Load Hypothesis (ILH)における認知処理水準の指標に基づき数値化した結果、より処理水準指標値の高いタスクを与えられた群(RM+TL Group)は、RM+MCQ Groupよりも語彙間の距離をより近く認識するようになる傾向がみられた(図1)。本研究により、既習語における語彙ネットワーク構造の変化を促すためには、深い意味的処理が必要であるという知見が得られた。多読後に既習語の語彙ネットワークがどのように変化したかについての可視化データを蓄積することにより、学習者の語彙知識発達についてのポートフォリオを構築し、フィードバックすることが可能であると考えられる。

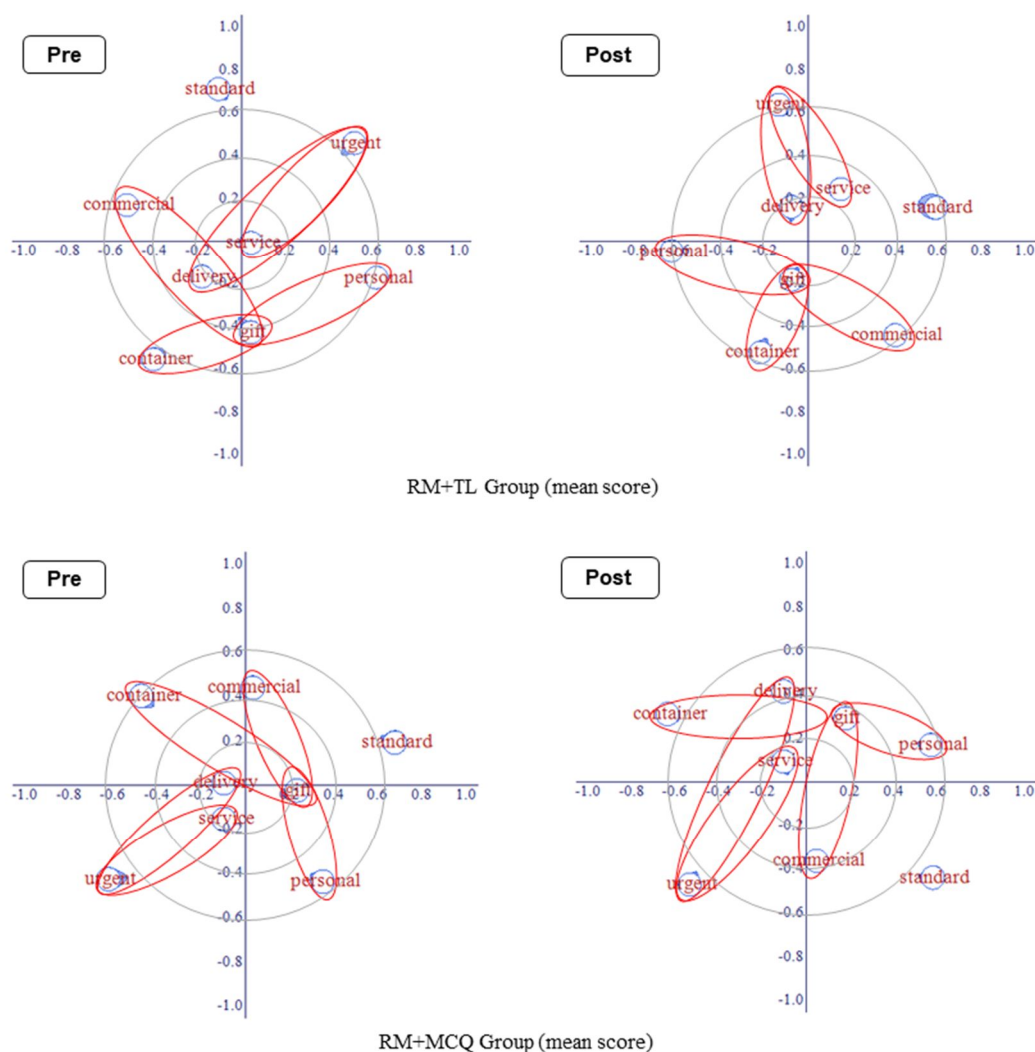


図1. AMISESCAL mapping of each group before (Pre) and after (Post) the task

(2) ネットワーク構造の可視化による語彙知識の構造化支援を目的とした本研究プロジェクトの基礎研究として、ES(A)Pの観点からの分野固有語彙ネットワークの拡張・深化過程の可視化を試みた。まず、情報科学と哲学という異なる分野の学術論文に内在する英語語彙の共起関係をアソシエーション分析により抽出し、共起関係に基づく語彙のネットワーク構造の視覚化を行った。その結果、どの論文でもキーワードを中心とした複数の小規模ネットワークが相互に独立して全体を構成しているという共通点が認められた。一方、情報科学分野では高頻度機能語とキーワードを含む固定フレーズが多用されるのに対し、哲学分野ではより内容に則してキーワードが共起していることが示され、ES(A)P教育における語彙指導に関し新たな示唆をもつ知見が得られた。その知見を基に、異分野で使用される表現の差異を描出することを目的として、データマイニングの一手法であるアソシエーション分析を援用し、さらに抽出されたアソシエーションルールをネットワーク図として視覚化した。情報科学分野の論文と医学分野の症例報告からそれぞれ約10万語を収集し、Rパッケージ apriori を用いて一文中に共起する語の組み合わせをアソシエーションルールとして抽出した。その結果、前者では4,726文から61個の共起パターンが、後者では4,848文から105個の共起パターンが抽出された。次に、ネットワーク視覚化ツール Gephi を用いて共起パターン相互の関連性を把握した(図2)。分析結果は、分野間の違いに加えて論文と症例報告というジャンル間の差異を明確に示しており、アソシエーション分析とネットワーク分析の組み合わせによる特徴描出が有効な手段であることが示された。また、共起パターンの視覚化がESPライティングに果たす役割も示唆された。

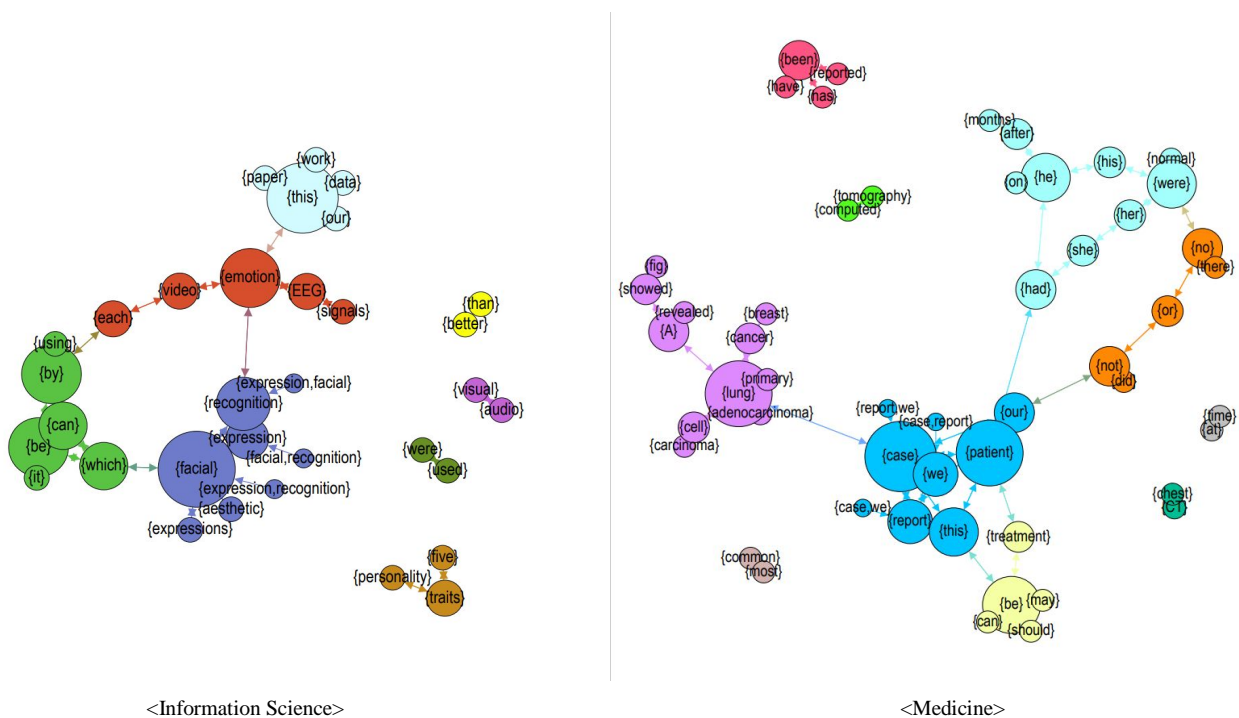


図2. Lexical network detected from papers and reports in different academic disciplines

(3) 本研究課題では英語語彙使用のあり方を視覚化し、テキストのジャンルや学術分野、および、書き手の属性(母語, 熟達度, 学術的背景等)による相違を把握することを目指して複数の分析手法を援用してきた。そのうちのひとつが、データマイニングの一手法であるアソシエーション分析を利用し、英文の文単位での単語間の繋がりを解析し、テキストに潜在する語彙ネットワークを視覚化する分析手法である。この手法では、大きく分けて、前処理として英語で書かれた文章の構造化を行った上で、アソシエーション分析を実行し、得られた共起関係を元にネットワーク図を描画する、といった段階を経る。本研究では、この分析手法を自動化することで、大量のデータについて各分析ツールへの習熟を必要とせず実行できる「英語語彙ネットワーク構造可視化システム」(LexNetViz)の開発に取り組んだ。

本システムではウェブブラウザ上で対話的にパラメータを設定しながら以下のプロセスが実行される。

テキストデータの前処理

非構造化データである英語文書を 用いるアソシエーション分析で利用するトランスアクションデータとして読み込み可能な形に整形し、テキスト形式で書き出す。分析対象となるテキストデータを読み込んだ時点で単語単位の頻度が集計され、上位20語がプロットされる。必要であればこの頻度プロットを繰り返し、削除対象語を確認した上で次の段階に進む。

アソシエーション分析の実行

で得られたテキストデータを元に R の [arules] [arulesviz] パッケージを利用してアソシエーション分析を実行する。その際、信頼度等の閾値を対話的に調整し、抽出される共起パターン数を確認しながら設定することで、必要な信頼度を持つ共起パターンを抽出する。

ネットワーク図の描画

のアウトプットである共起パターンをネットワーク図として描画し、png 形式で書き出す。また、共起パターンの相互関係は gml 形式でも書き出され、gephi など、他のネットワーク分析アプリケーションでの利用も可能とする。

本システムの開発により、語彙ネットワークの拡張・深化過程を可視化する新しい手法が提供された。

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Aotani N. and Takahashi S.	4. 巻 -
2. 論文標題 An Analysis of Japanese EFL Learners' Lexical Network Changes.	5. 発行年 2021年
3. 雑誌名 Front. Educ. 5:621437.	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.3389/feduc.2020.621437	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計4件（うち招待講演 0件/うち国際学会 4件）

1. 発表者名 Sugino, N., Aotani, N., Koga, Y., Shojima, K., and Fraser, S.
2. 発表標題 Disciplinary Differences in Lexical Co-occurrence Patterns.
3. 学会等名 ALAA2018（国際学会）
4. 発表年 2018年

1. 発表者名 Aotani, N., Sugino, N., Koga, Y., Fraser, S., and Shojima, K.
2. 発表標題 A Preliminary Analysis of Learners' Lexical Network Changes.
3. 学会等名 HIC2019（国際学会）
4. 発表年 2019年

1. 発表者名 SUGINO Naoki, AOTANI Noriko, KOGA Yuya, SHOJIMA Kojiro
2. 発表標題 Vocabulary demands of academic papers from the structural perspective.
3. 学会等名 CAES International Conference: Faces of English 2（国際学会）
4. 発表年 2017年

1. 発表者名 AOTANI Noriko, SUGINO Naoki, FRASER Simon, KOGA Yuya, SHOJIMA Kojiro
2. 発表標題 A Comparison Between L2 Learners' Conceptions of Their Vocabulary Network and Visualized Representations obtained from a Directional Statistics Model, AMISESCAL.
3. 学会等名 The Applied Linguistics Conference (ALANZ / ALAA / ALTAANZ) (国際学会)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	杉野 直樹 (Sugino Naoki) (30235890)	立命館大学・情報理工学部・教授 (34315)	
研究分担者	荘島 宏二郎 (Shojima Kojiro) (50360706)	独立行政法人大学入試センター・研究開発部・准教授 (82616)	
研究分担者	Fraser Simon (Fraser Simon) (10403510)	広島大学・外国語教育研究センター・教授 (15401)	

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	古賀 友也 (Koga Yuya)		

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	高橋 晋也 (Takahashi Shin'ya) (70260586)	東海学園大学・心理学部・教授 (33929)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関