

令和 5 年 6 月 6 日現在

機関番号：32612

研究種目：挑戦的研究（萌芽）

研究期間：2020～2022

課題番号：20K21831

研究課題名（和文）脳損傷と加齢が潜在意味構造に与える影響の解明：自然言語処理に基づいたアプローチ

研究課題名（英文）Natural language processing approach for the effects of brain damages and aging on latent semantic structure

研究代表者

板口 典弘（Itaguchi, Yoshihiro）

慶應義塾大学・文学部（三田）・助教

研究者番号：50706637

交付決定額（研究期間全体）：（直接経費） 5,000,000円

研究成果の概要（和文）：本研究では、自然言語解析に基づいた意味ネットワーク解析手法を開発し、脳損傷（失語症状）および加齢が、ネットワーク構造に与える影響を解明することを目指した。研究の結果、以下の成果を得た。まず、日本語およびノルウェー語Wikipediaを対象としたLDA解析による、動物単語に関する定量的な意味構造を得た。次に、日本人・ノルウェー人を対象とした言語流暢性課題の成績を、それぞれの意味構造に基づいた視覚化方法を開発した。さらに、両集団において、単語産出インターバルと意味連想の強い関連性を得た。最後に、患者を対象とした研究によって、加齢や脳損傷・認知機能低下が意味構造変容に与える影響を定量化した。

研究成果の学術的意義や社会的意義

本研究では、心理学的手法と情報学的手法を統合し、私たちの脳内において、「意味」がどのように表現されているかを検討した。この検討によって、加齢や脳損傷、認知症によってどのように個々の意味構造が変容するかについて、その一端を明らかにすることができた。また、日本人とノルウェー人を対象とした実験を行った結果、文化によっても、意味構造が大きく異なる可能性を示した。これらの成果は、今後の何らかの認知機能が低下した方への的確な理解や、効率的なリハビリテーションの提案に繋がる可能性がある。

研究成果の概要（英文）：We aimed to develop a semantic network analysis method based on natural language analysis and to elucidate the effects of brain damage and aging on the network structure. First, quantitative semantic structures of animal words were obtained from LDA analyses of Japanese and Norwegian Wikipedia. Second, we developed a visualization method to interpret the performance of Japanese and Norwegian subjects in a verbal fluency task based on their semantic structures. Furthermore, we found a strong relationship between word production interval and semantic association in both populations. Finally, we quantified the effects of aging, brain damage, and cognitive decline on semantic structure change through a patient study.

研究分野：心理学

キーワード：心理学 言語 認知症 失語症 自然言語処理

## 1. 研究開始当初の背景

ヒトがその脳内に持つ“意味”は、心理学だけでなく、多くの学術分野において研究対象となっている。本研究では特に、脳損傷と加齢がヒトの持つ意味構造に与える影響の解明を目指した。

意味を定量化しようとする試みは心理学モデルを用いて広範なアプローチがおこなわれてきた。近年注目されているアプローチが自然言語処理を用いたアプローチである。ただし、国際的には、自然言語解析に基づいた意味構造へのアプローチが散見される(Pakhomov et al. 2014, 2015; Woods et al. 2017)ものの、わが国においては、加齢や脳損傷が引き起こす意味構造変化に対する同様のアプローチは未だ存在しなかった。この点で、日本語を対象とした意味構造研究は萌芽以前の状態であった。また、解析手法については、海外研究の多くがLSA(潜在意味分析)を用いているのに対して、本研究はLDA(潜在ディリクレ配分法)を用いた点で新規性と強みがあった。LDAは一般化可能性の強いものの、データの解釈が困難であるという特性を持っていたのに対し、本研究では心理学モデルおよび時間解析を適用しその解釈困難性を解決した。

また、本研究は応用的意義も強く持っていた。脳損傷患者や認知症患者を対象とする臨床現場では患者の認知機能を評価するために、日常的に様々な検査が実施されている。その代表例がVFT(言語流暢性課題)である。通常、VFTでは一定時間内に回答できた単語の総反応数が認知能力の指標として扱われる。一方で、その回答内容、すなわち単語の意味特性や時系列的側面は、患者の意味構造を反映するものとして捉えられてきたものの、定量的・客観的に扱う方法がないために十分に検討することができなかった。我々は、自然言語処理に基づいた標準的な意味ネットワークモデルを構築することによって、この問題を解決できると考えた。

## 2. 研究の目的

本研究では、自然言語解析に基づいた意味ネットワーク解析手法を開発し、脳損傷(失語症状)および加齢が、ネットワーク構造に与える影響を解明することを目指した。具体的には、(1)LDAを用いた日本語における単語間の客観的な意味連想関係を抽出・数値化、(2)個人の意味ネットワーク構造の特徴を定量化する方法論の確立、(3)脳損傷・加齢による意味構造変容の定量化、の3点を下位目標とした。

## 3. 研究の方法

(研究A)日本語データベースのLDA解析およびそれに基づいた健常者VFTの成績解析

大学生(日本語話者)49名を対象とし、動物語を用いたVFTを実施した。VFTの回答に反映される個人の意味ネットワークを抽出するために、日本語版Wikipediaの記事に対して、LDA(潜在ディリクレ配分法)を用いたトピック分析(Blei, 2003)を実施した。これによって、動物というカテゴリにおける意味構造、すなわちある単語が出現した際にあるトピック(サブカテゴリ)が存在する確率を得ることができる。本研究ではこれをトピック確率と呼ぶ。サブカテゴリに対するトピック確率を用いて、VFTにおける単語クラスタリングを客観化・定量化した。さらに、得られた意味クラスタ情報に加えて、時間情報・単語頻度情報を統合した視覚化方法を提案する。研究Aは目的(1)と(2)に相当する。本成果は現在論文執筆中である。

(研究B)失語症患者および前頭葉機能障害患者に対してVFTを行い、リハビリテーションに伴う各種指標の変化を検討した。研究Bは目的(3)に相当する。本成果は大門ら(2021)にて発表済みである。

(研究C)ノルウェー語データベースのLDA解析およびそれに基づいた健常者・高齢者・認知症患者のVFT成績解析

ノルウェー語話者計66名を対象とし、動物語を用いたVFTを実施した。研究Aと同様の方法を用いて、VFT成績の定量化を行った。さらに、認知症患者においてはエラーの意味・時間解析もおこなった。研究Cは発展的研究である。

## 4. 研究成果

(研究A)

まず、LDAベースの意味情報・時間情報・単語頻度情報を統合した視覚表現方法を確立した。この一例をFigure 1aに示す。次に、LDAによって推定された意味クラスタリングの妥当性を、時間情報(単語産出インターバル)を用いて検証した。その結果、クラスタ内インターバルの方がクラスタ外インターバルよりも有意に短いことが明らかとなった(Figure 1b)。この結果は、LDAに基づいたクラスタリングでも(LDA-based)、検査者がクラスタリングした場合でも(Human-based)同様であった。また、Mean Cluster size(MCS)はLDA-basedの方がHuman-basedよりも大きくなった(Figure 1c)。これらの結果は、LDA-based clusteringでもHuman-basedと同様に意味関連情報を反映できていること、さらにHuman-basedでは検知できないレベルの意味的関連もClusteringできていたことを示唆する。

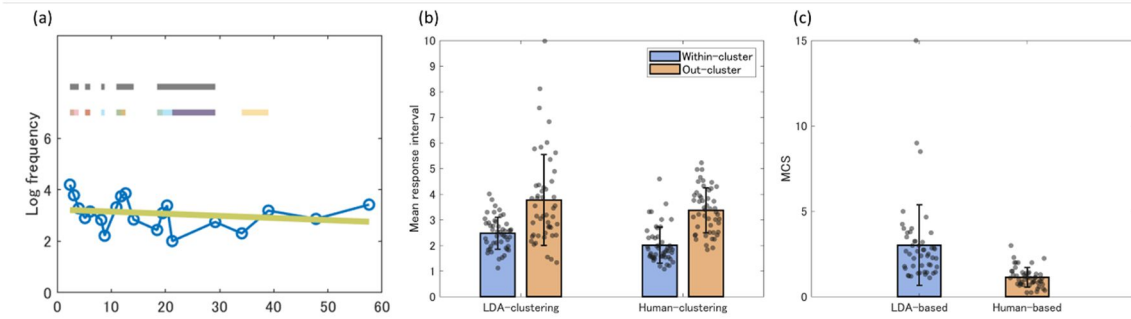


Figure 1. 日本人健常若年者を対象とした VFT 成績解析

(研究 B)

日本人失語症患者および前頭葉機能障害患者に対して言語流暢性課題の連続実施をおこない、その継時的変化を明らかにした。損傷部位の異なる複数の患者に課題実施をおこなったが、ここでは特に、大門ら (2021) の結果を報告する。

本研究では、前頭葉機能に低下を示した脳血管疾患 2 症例に対して、継時的に語想起課題を実施し (週 5 回を 4 週あるいは 5 週)、産出単語の時間情報と頻度情報を併せた検討を行った。両症例とも、総反応数は増加していったものの、評価期間後期で横ばいとなった (Figure 2A)。各回の単語出現頻度の平均値は大きく変化しなかったが、コンスタントに比較的低頻度な新規単語が出現した (Figure 2E, F)。また、長い時間スケール (評価期間) および短い時間 (1 分間の課題内) スケールそれぞれの時間スケールにおいて、高頻度語から低頻度語に移行するという単語頻度効果が確認された (Figure 2G)。さらに、症例 1 名においては、課題内の頻度効果が、評価期間を通じて強まっていった。本研究で観察された 2 つの時間スケールにおける頻度効果およびその変化を、言語流暢性課題内における言語性要因と実行機能要因のダイナミズム、および実行機能の回復という観点から検討した。

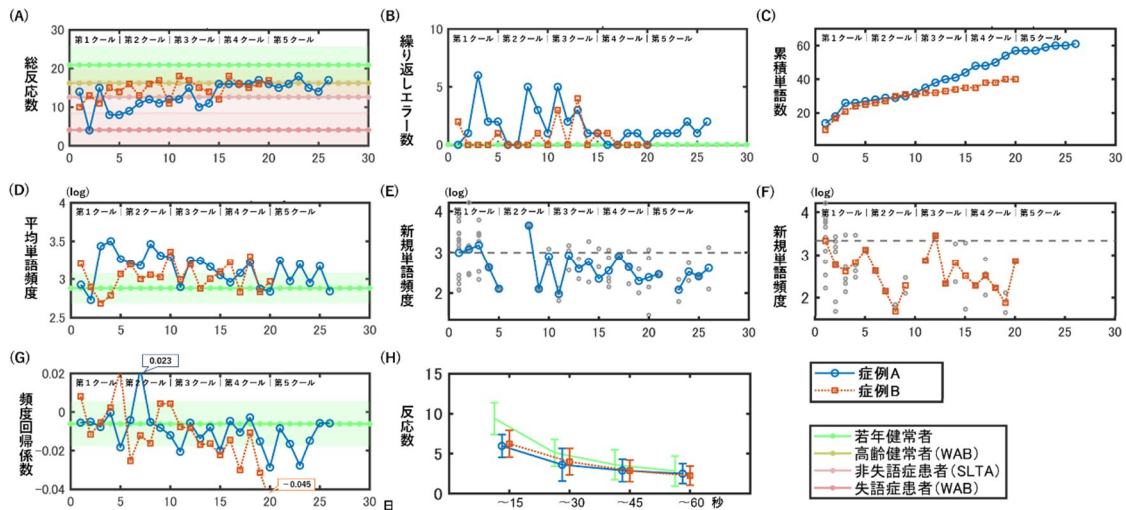


Figure 2. 日本人脳損傷患者を対象とした VFT 成績解析

(研究 C)

日本語を用いた解析と同様に、ノルウェー語版 Wikipedia をデータベースとした LDA 解析をおこない、さらにそのトピック寄与率をもとに、アルツハイマー病患者、高齢健常者および若年健常者の言語流暢性課題成績を解釈した。具体的には、言語流暢性課題のエラーがどのような意味ネットワーク間の連想関係に基づいて、どのようなタイミングで出現するかを定量的に明らかにした (Figure 3)。本成果は Itaguchi et al. (2022) によって発表された。さらにこの知見を補強するために、エラー以外の要素についての解析をおこなった。その結果、日本人健常者によって得られた意味的関連性と時間的要素の強い結びつきを、ノルウェー人を対象にしたデータによっても再現した。この成果は現在論文執筆中である。

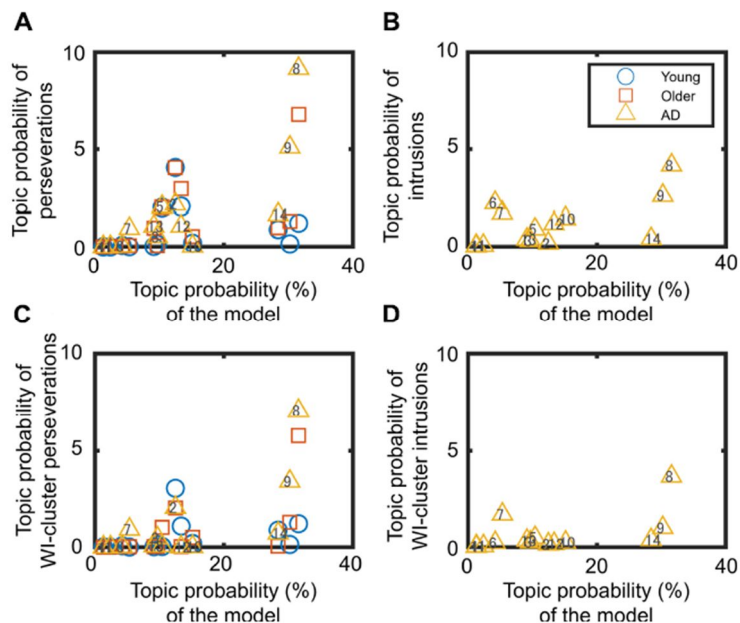


Figure 3. ノルウェー人脳損傷患者を対象とした VFT 成績解析

その他，意味構造の文化差，日本人高齢者を対象とした VFT 成績の解析，日本人失語症患者を対象とした VFT 成績の解析など，多くの知見を得た。これらは順次論文執筆用である。また，日本語・ノルウェー語 Wikipedia を対象とした動物単語に関する意味・頻度データベースも公開予定である。

## 5 . 引用文献

1. Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.* 3, 993–1022.
2. 大門正太郎, 能登栞, 板口典弘. (2021) 繰り返し実施した言語流暢性課題成績に対する時間情報と頻度情報を用いた定量的解析. *高次脳機能研究*, 41(4), 387-396.
3. Itaguchi Y., Castro-Chavira SA, Waterloo K, Johnsen SH, Rodríguez-Aranda C. (2022) Evaluation of error production in animal fluency and its relationship to frontal tracts in normal aging and mild Alzheimer's disease: A combined LDA and time-course analysis investigation. *Frontiers in Aging Neuroscience*, 13, 710938.
4. Pakhomov, S. V. S., Eberly, L. E., and Knopman, D. S. (2018). Recurrent perseverations on semantic verbal fluency tasks as an early marker of cognitive impairment. *J. Clin. Exp. Neuropsychol.* 40, 832–840.
5. Woods, S. P., Conover, E., Rippeth, J. D., Carey, C. L., Gonzalez, R., Marcotte, T. D., et al. (2004). Qualitative aspects of verbal fluency in HIV-associated dementia: a deficit in rule-guided lexical-semantic search processes? *Neuropsychologia* 42, 801–809.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 1件/うち国際共著 1件/うちオープンアクセス 1件）

1. 著者名 Itaguchi Y., Castro-Chavira SA, Waterloo K, Johnsen SH, Rodriguez-Aranda C.	4. 巻 13
2. 論文標題 Evaluation of error production in animal fluency and its relationship to frontal tracts in normal aging and mild Alzheimer's disease: A combined LDA and time-course analysis investigation	5. 発行年 2022年
3. 雑誌名 Frontiers in Aging Neuroscience	6. 最初と最後の頁 710938
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/s00221-022-06350-6	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 該当する

1. 著者名 大門正太郎, 能登栞, 板口典弘	4. 巻 41
2. 論文標題 繰り返し実施した言語流暢性課題成績に対する時間情報と頻度情報を用いた定量的解析	5. 発行年 2021年
3. 雑誌名 高次脳機能研究	6. 最初と最後の頁 387-396
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計3件（うち招待講演 0件/うち国際学会 0件）

1. 発表者名 大門正太郎, 能登栞, 板口典弘.
2. 発表標題 語想起課題の継時的評価における産出単語の時間情報・頻度情報を用いた解析.
3. 学会等名 44回日本高次脳機能障害学会学術総会
4. 発表年 2020年

1. 発表者名 大門正太郎, 板口典弘.
2. 発表標題 失語症例に対する動物単語の語想起訓練と視覚性呼称訓練の相互作用
3. 学会等名 第46回日本神経心理学会学術集会
4. 発表年 2022年

1. 発表者名 大門正太郎, 板口典弘.
2. 発表標題 視覚性呼称と語想起課題の成績が併行して上昇した失語症例 - 産出単語の頻度情報を用いた解析 -
3. 学会等名 第45回日本高次脳機能障害学会学術総会
4. 発表年 2021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	吉原 将大  (Yoshihara Masahiro)  (70822956)	早稲田大学・文学学術院・その他(招聘研究員)   (32689)	VFT課題実施, VFT課題解析, データベース構築, 論文執筆

	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	大門 正太郎  (Daimon Shotaro)		
研究協力者	クラウディア ロドリゲス-アラランダ  (Rodriguez-Aranda Claudia)		

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
---------	---------

ノルウェー	Tromso University			
-------	-------------------	--	--	--