

令和 6 年 5 月 29 日現在

機関番号：34426

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K00491

研究課題名（和文）グラフ理論を用いた自然言語システムの構造構築メカニズムの本質と存在理由の考察

研究課題名（英文）Essence and ontology of structure-building mechanisms of natural language system: a graph-theoretical approach

研究代表者

有川 康二 (Arikawa, Koji)

桃山学院大学・国際教養学部・教授

研究者番号：80299023

交付決定額（研究期間全体）：（直接経費） 1,600,000円

研究成果の概要（和文）：三つのテーマで英語論文を出版した（(a)は査読付きの国際学術雑誌）。(a) ホモ・サピエンス（私たち）の脳内の言語システムが産み出す文構造内で働くバランス力（量子力学（物理学）と同じ数学的構造を持つ行列統辞論の枠組み）、(b) 言語システムの誤差（エラー）最小化システムとしての進化（生物学的な免疫機構との類似性）、(c) グラフ理論（数学）の文構造への応用。言語システム研究は数学、物理学、生物学と統合可能である。

研究成果の学術的意義や社会的意義

鳥たちは飛びまくり、花たちは咲きまくり、犬たちは嗅ぎまくり、私たちは喋りまくる。人は言葉によって生きもすれば死にもする。人間は言葉の根本を何千年も探究したが解明できていない。言葉の根本の解明は人間社会にとって急務。時速11万キロで公転中の地球を含む宇宙の探究が急務であるように。言葉はヒト脳の言語システムの働き。言語システムは2個の要素を無限に結合する結合マシン。コミュニケーションに適応して進化したのではない。人間が騙し合い、殺し合う動物であるわけだ。なぜ言葉はこんな姿かたちなのか？なぜ私たちの脳だけがこんな言葉を産み出すのか？本研究はこの問題の解決への一歩である。

研究成果の概要（英文）：I published three papers regarding three topics ((a) was published in an international academic journal with review). (a) balance force working in sentential structures produced by the language system of Homo sapiens (us) (written in the matrix-syntactic framework, which shares the same mathematical structures with that of quantum mechanics (physics)), (b) language system has evolved as an error-minimization system (similar to the immune system in biology), and (c) application of graph theory (mathematics) to the language system. The human-language-system investigation can be integrated with mathematics, physics, and biology.

研究分野：生物言語学

キーワード：自然言語 ホモ・サピエンス脳 生成統辞論 グラフ理論 行列統辞論 文構造 誤差最小化 ウイルス・チェック

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

## 1. 研究開始当初の背景

本研究の背景として、国内外で本格的に始動した感のある「言語設計における第三要因 (物理法則)」を基に数理科学との将来的な統合を見据えた学際的研究に対して、本研究は如何なる貢献が可能かということがあった。本研究の出発点は、筆者を研究代表者とした科研費研究「統辞素性の診断法における数学的証明 (背理法) の有効性」(概略テーマ) である (奨励研究 (A); 1998-1999、及び、基盤研究 (C); 2001-2002, 詳細は後述)。言語システムの証拠提示の方法として背理法が多用される事実は、言語システムの公理と定理による体系化が可能であることを示唆する。そこから、言語と数学のアルゴリズムの共通点と相違点、言語計算の一般性と特殊性に関心を持ち、構造の性質を定量的に解析するグラフ理論、また、対称性を定量的に表現する群論を言語に応用し、その説明力を検証しながら論文を発表してきた。しかし、これらの考察では、「何故、自然言語はこのような構造を示すのか、何故、別の構造を示さないのか」という総合的な観点が希薄であった。本研究は、これらの科研費研究以降の考察を踏まえて問うに至った、この総合的な問いを探求することであった。具体的には、自然言語アルゴリズムに出現する構造へのグラフ理論の適用、自然言語システムの生成する構造的性質の定量化、変数消去のアルゴリズム、併合演算、MC が構造計算に要請するフェーズの本質と存在理由を考察する必要があった。

## 2. 研究の目的

自然言語システムの生成する構造に線型代数学の応用であるグラフ理論を適用し、言語構造の様々な性質の定量化を行い、それが内的併合やフェーズの存在理由、また、文法性判断の変動にどのように関わるかを考察し、具体的な仮説と経験的証拠を提示することが研究の目的であった。

## 3. 研究の方法

研究方法の根幹としては、線形代数学の代表的な応用分野であり、構造一般の探求に不可欠のグラフ理論を、言語システム研究に利用することであった。Chomsky (2013) が提案するような「統辞論におけるグラフ理論的探求の破棄」は早急であることを、グラフ理論応用の有効性を示すことで論駁することであった。また、自然界の複雑な構造を探求するための数学であるフラクタル幾何学も利用することも予定した。言語システムが生成する抽象構造を樹形図として近似した場合、フラクタル幾何学的複素空間に出現するデンドライト (樹状構造) との相同性はあるのか。言語構造の本質の解明に向け、経験的な証拠を補強する為に数学的な道具 (グラフ理論とフラクタル幾何学) を使用することを掲げた。

本研究は、自然言語システムのアルゴリズムが生成する構造を、数学という人工言語のアルゴリズム (ここではグラフ理論やフラクタル幾何学) の鏡に映し出した時に何が見えるかを検証するものであった。何が自然言語の特性で、何が人工言語の特性なのか、自然言語の構造の本質は何か、内的併合とフェーズの本質と存在理由は何かに関する仮説を提出する目標を掲げた。統辞論と、数学や物理学に代表される基礎科学との将来的な連携に向けた基礎的な研究に僅かでも

貢献することを標榜した。

#### 4. 研究成果

##### 論文

当該期間に以下の論文を出版した。論文(a)は査読付きの国際研究ジャーナルであり、論文(b)、(c)は本学の紀要である。

- (a) Arikawa, K. (2022) Internal merge: Why does it work this way? A matrix syntactic approach to argument chain. *Open Journal of Modern Linguistics*, 12-3, 336-365.
- (b) Arikawa, K. (2023a) Non-referentialist  $C_{HL}$  as error minimization: Toward a valuation-free Agree model. *St. Andrew's University Bulletin of the Research Institute*, 48-3, 79-104.
- (c) Arikawa, K. (2023b) Putting old wine<sup>ECP, connectedness condition, island</sup> into new wineskin<sup>graph theory</sup>. *St. Andrew's University Bulletin of the Research Institute*, 49-1, 119-158.

論文(a)は行列統辞論の枠組みを利用し、内的結合によって生成される文構造内の均衡力を計算、分析、考察したものである。その結果、内的結合により形成される文構造には右下の図(Figure 6)のような均衡が潜在していることが判明した。構造の背骨部分(チェック模様)は安定しており、構造形成部分はエネルギー拡大傾向、内的結合のコピーは縮小傾向にあることが分かった。行列統辞論(matrix syntax)とは、Orús Roman, Roger Martin & Juan Uriagereka.

2017. *Mathematical foundations of matrix syntax*. arXiv.org, Cornell University. や、Martin Roger, Román Orús & Juan Uriagereka. 2019. *Towards matrix syntax*. In *Catalan Journal of Linguistics: Special Issue*, 27-44 等の論文で提案されたヒト言語システム探究の可能なアプローチ

で、量子力学の数学的構造を言語システムの内的結合分析に応用しようとするもので、線形代数学の行列計算と利用する。将来的な数回のブレークスルーを経験した物理学との統合を見据えたアプローチを提案したものである。これらの論文は、特に内的結合により形成される連鎖(chain)の数学的構造は、量子力学の根幹である「重ね合わせ」(superposition)のそれを相同であることを主張している。重ね合わせの現象は、光子(フォトン)等の素粒子の本質的な特徴であり、脳が量子コンピューターを実現している自然物である可能性を考えると、将来の脳研究の重要な基礎研究の一端を担うものと考えられる。

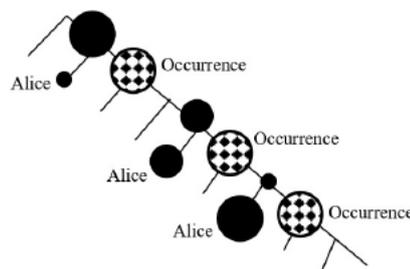


Figure 6. Geometrical image of the second sub-step of chain formation (modified).

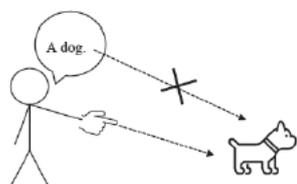


Figure 3: We refer but words do not

論文(b)は「語は物を指し示す」という指示主義というドグマ(右のFigure 3参照)の否定の根拠を示しつつ、ヒト言語システムが誤差(エラー)最小化システムであることを論じた。「語は物を指示する」という指示主義(referentialism)は、いまだにプロの哲学者の間でも根深いドグマの一つであり、ヒト言語

研究の根幹が未解明であることを示唆する問題である。

論文(c)は文構造にグラフ理論を応用し、構造生成異常の数値化により、文法性判断の根拠を定量的に分析、考察していける可能性を示した。例えば、以下の相対的に文法的な(a)、非文法的な(b)の文構造には各々下の Figure 5 と Figure 6 で示すような情報の流れが潜在していることが分かった。

(a)? a person who they spoke to e because they admired e

(b)\* a person who they spoke to e because e admired them

非文法的な(b)の文構造では従属節の主要部投射で情報の流れが遮断されていることが分かった。(a), (b)の差は Chomsky (1981)の Empty Category Principle (ECP; 空範疇原理)の不備を指摘する為に Kayne

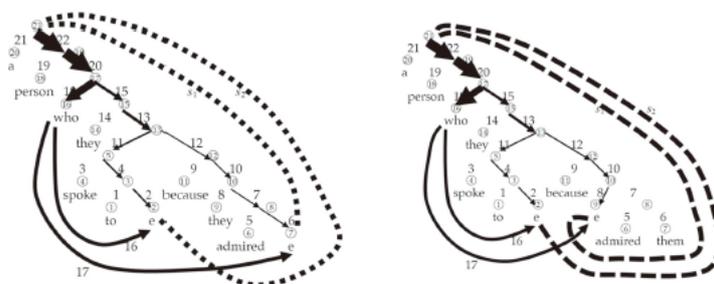


Figure 5: Force balance hidden in Figure 4 Figure 6: Force balance hidden in Figure 8

(1983)が Connectedness Condition (CC; 結合性条件)を提案する際に使用したものである。奇しくも CC はグラフ理論との統合を目論んでいた。グラフ理論とは線型代数学の巨大の応用分野の一つで、具体、抽象問わず、いわゆるネットワーク (情報の流れを点(nodes)と線(edges)で結合したもの)の均衡を数量的に計算するツールである。ヒト脳は、文構造という無限のグラフ構造(内的結合によりループを内包している)を生成する臓器であり、そのような文グラフの数学的性質についてグラフ理論を利用して炙り出していくことは今後の統辞研究に対して新鮮な知見を提供できる。

## 著作

また、継続して以下の著書の出版に向けて鋭意努力している。現在、Springer Nature 社による外部査読の結果待ちの状況である。

Arikawa, K. In review. *Elements of generative syntax: How rationality tackles the language system* (邦訳: 生成統辞論の基礎: 人間の合理性は言語システムの説明に如何に立ち向かっているか), 約359頁, 使用言語: 英語。

当該書籍は、過去20年以上にわたり著者が本学で「日本語学概論」「日本語文法論」「Japanese Studies」(現Global Experience Seminar: Exploring the human language system with overseas students[グローバル体験セミナー: 交換留学生たちと一緒に言語理論を学ぼう]); 英語による授業)の授業の中で個別部分的に教科書として利用しつつ学生たちと切磋琢磨しながら継続して準備したものである。従って、過去の学生たちへの学恩に報いるためにも当該書籍は是非出版したいと思っている。例えば、現在、Global Experience Seminar の受講生にマサイ語(ナイル・サハラ語族、基本語順: VSO)を母語とする交換留学生がいるが、その学生にマサイ語の文データを提供してもらい、マサイ語の基本的文構造が英語や日本語の文構造と同じであるというような授業内容を構成することが可能となっている。地球上の全ての言語は同じホモ・サピエンス語の方言であることを受講生達に感じて欲しいと願っている。

当該書籍は、人間が数千年以上かけて挑み続けている、ホモ・サピエンス脳内に存在する自然言語計算システムの解明という難問に対して、現時点で人間が自らの理性を駆使してどのように取り組んでいるか的一端（生物言語学の枠組み）を示したものである。例えば、以下のような論点を掲げている。

- スイスの言語学者フェルディナン・ド・ソシュール（1857-1913）は一般言語学を代数学や幾何学のような公理体系として再編しようとしていた。
- アメリカの言語学者ノーム・チョムスキー（1928-）は言語を物理学、生物学と統合しようとしている。
- 言葉に関するドグマとその問題点を洗い出す。例えば、以下のようなドグマを再考する：言葉は単純である、言語は自然物ではない、言語システムはコミュニケーションに適応して進化した、等。
- ヒト以外の脳では、知覚・運動システムと概念・志向システムが直結しているが、ヒト脳では、それらが言語システムを介して間接的に繋がっている。その為にヒトは嘘をつく動物になっている。
- 言語システムの根幹的な操作である結合の特性は何か？
- 言語システムの基本性質とは何か？
- 言語システムにとっての第三要因である「最小計算の法則」とは何か？
- 言語システムは偽装ウイルス・チェック・システムである。
- 他動詞とは何か？
- 主語とは何か？発音されないコピーが主語であるとはどういうことか？
- 濁音化、連濁、ライマンの法則の本質は何か？
- 言語システムが生成するグラフ構造（情報が流れる点と線のネットワーク）と集合理論を基盤とする結合操作をもとに言語システムを分析、考察するという作業を行なっている。
- 言語システム分析には数学の交換法則、結合法則に関連するテーマ設定が可能である。

データとして20種類以上の言語を利用し、言語システム研究の数学、物理学、生物学との将来の統合を見据えた論点を提供している。

5. 主な発表論文等

〔雑誌論文〕 計3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 3件）

1. 著者名 Koji Arikawa	4. 巻 12
2. 論文標題 Internal merge: Why does it work this way? A matrix syntactic approach to argument chain	5. 発行年 2022年
3. 雑誌名 Open Journal of Modern Linguistics	6. 最初と最後の頁 336-365
掲載論文のDOI（デジタルオブジェクト識別子） 10.4236/ojml.2022.123025	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 -

1. 著者名 Koji Arikawa	4. 巻 48
2. 論文標題 Non-referentialist CHL as error minimization: Toward a valuation-free Agree model	5. 発行年 2023年
3. 雑誌名 St. Andrew's University Bulletin of the Research Institute	6. 最初と最後の頁 79-104
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 オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Koji Arikawa	4. 巻 49
2. 論文標題 Putting old wine (ECP, connectedness condition, island) into new wineskin (graph theory)	5. 発行年 2023年
3. 雑誌名 St. Andrew's University Bulletin of the Research Institute	6. 最初と最後の頁 119-158
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 オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計0件

〔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
---------------------------	-----------------------	----

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