【学術変革領域研究（B）】

大規模言語モデルを駆使したインシリコ言語学の創成

研究代表者	東京大学・総合文化研究科・言語情報科学専攻・准教授 川崎　義史（かわさき　よしふみ）　研究者番号：40794756
研究課題情報	課題番号：26B101　研究期間：2026年度～2028年度 キーワード：大規模言語モデル、言語学、自然言語処理、神経科学

なぜこの研究を行おうと思ったのか（研究の背景・目的）

●研究の背景

2022年以降、ChatGPTに代表される**大規模言語モデル（Large Language Model; LLM）**は日進月歩の勢いで進化している。LLMの驚異的な言語理解能力・生成能力を目の当たりにし、LLMを活用した新たな言語研究「**インシリコ言語学**」の構想に至った（図1）。

従来の言語研究は、理論や内省（専門用語でin papyro ラテン語で「紙の上で」の意）、自然な発話やテキストの収集（in situ「その場で」）、統制実験（in vitro「試験管内で」）、あるいは生体実験（in vivo「生体内で」）により得られたデータに基づき展開されてきた。これらの手法の有効性は論を俟たない。しかし、その一方で、従来手法では直接的な観測や定量化が困難な現象（例：言語変化、意味）や、物理的・倫理的制約により実施が困難な実験（例：反実仮想実験、介入実験）も存在する。そこで、本研究では、LLMを駆使したインシリコな手法により、従来手法を補完する形で、言語研究を加速化させることを目指す。

＊インシリコ（in silico）とは？

「インシリコ」は、ラテン語で「シリコン上で」を意味する。コンピュータの半導体にシリコンが使用されていることに由来する表現で、**コンピュータによるシミュレーションやデータ解析**に基づく手法を指す。

インシリコな手法は、例えば、タンパク質の構造予測を行う人工知能AlphaFoldに代表されるように、化学・生物学・創薬等の分野でも大きな注目を集めている。実世界のタンパク質を直に対象とする代わりに、人工知能の中に再現した「タンパク質」を対象とするのである。インシリコな手法により、膨大なコストを要する試験管内（in vitro）実験を効率化し、その成果を実世界に還流することで、研究サイクルの加速化が起きている。

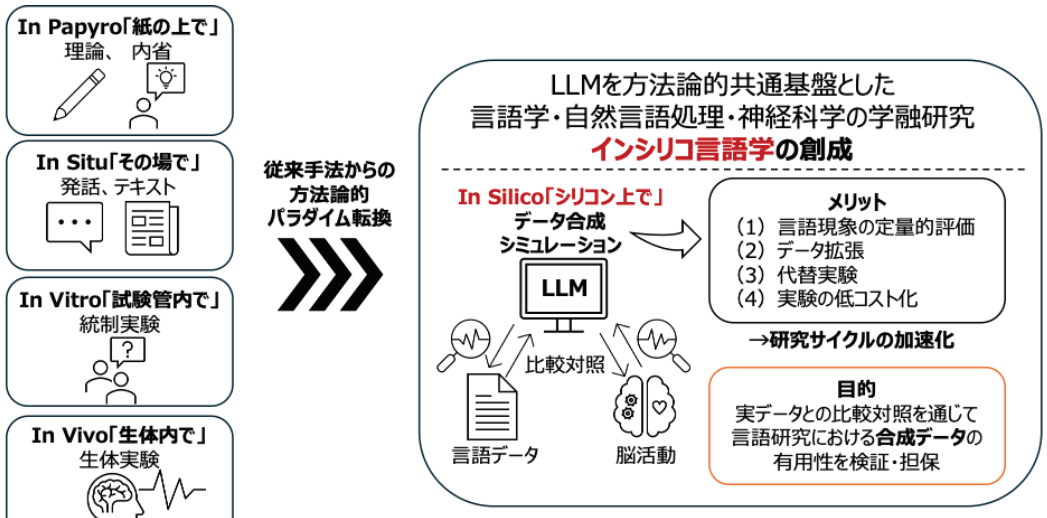

図1　インシリコ言語学の全体像

●研究の目的

本研究の目的は、言語研究における合成データ（テキストなどLLMが生成するデータ全般）の有用性を検証・担保することである。具体的には、合成データが実データ（テキスト、言語学的知見、脳活動情報など）をどの程度再現できているか検証する。そのために、LLMを共通のツールとして、**言語学・自然言語処理・神経科学**の知見や手法を総動員する。

現時点では、合成データを人間言語の研究に利用することに対して懐疑的な意見もある（例：LLMの生成するテキストは、人間が書いたものと性質が異なるのでは？）。しかし、もし合成データが実データを一定程度再現できるのであれば、合成データを言語研究に活用する道が開ける。同一ではないものの何らかの観点で同一視できるからこそ人間の代わりに動物で代替実験が行われるように、収集が高コストな実データの代わりに低コストな合成データを利用することで、言語研究の加速化が期待される。

インシリコ言語学とは、いわば、LLM上に再現した「言語」を介した言語研究である。その真骨頂は、LLMによる**データ合成**と**シミュレーション**である。これにより、（1）意味や認知など従来手法では評価が困難だった言語現象の定量的評価、（2）欠損値補完などのデータ拡張、（3）物理的・倫理的制約等により実施が困難な実験の代替、（4）被験者を要する高コストな実験の低コスト化、などの可能性が広がる。

この研究によって何をどこまで明らかにしようとしているのか

A01 コーパス合成学

LLMを利用して生成した言語データ（合成コーパス）が、実コーパスをどの程度近似できているか解明することを目的とする。また、人間が有するとされている内的概念（例：事態の捉え方）をLLMが保持しているかどうか、言語の生成過程や学習過程におけるLLMと人間の共通点・相違点は何なのか検証する。これにより、実コーパスの代替としての合成コーパスの有効性を探究する。

A02 内在特性測定学

単語の印象や意味など、言語データの表層には顕現しない特性（内在特性）を測定するための手法の創出を目的とする。いわば、言語データ用の顕微鏡やレントゲンのようなツールを創り出すことに取り組む。内在特性は、LLMの内部表現（ベクトル）に各種演算を適用することで測定する。また、LLMを被験者とみなした内在特性の測定手法の開発も目指す。

A03 言語シミュレーション学

言語研究におけるシミュレータとしてのLLMの有効性と限界を検証することを目的とする。具体的には、直接的な観測が困難な言語現象（例：言語変化、言語接触）のモデル化や、物理的制約や高コストのために実施が困難な実験（例：多言語の話者を要する通言語的実験）の代替として、LLMがどのように、どの程度、活用できるのか、その可能性を探究する。

A04 比較神経表現学

LLMを単なる言語処理モデルとしてではなく、ヒト脳における意味情報表現を理解するためのモデルとして捉え、その有効性と限界を検証することを目的とする。知覚・認識・解釈・言語化に至る一連の情報処理過程を対象に、脳活動とLLM内部表現の関係を体系的に調べることで、言語と非言語をまたぐ汎用的な意味情報表現のあり方を明らかにすることを目指す。

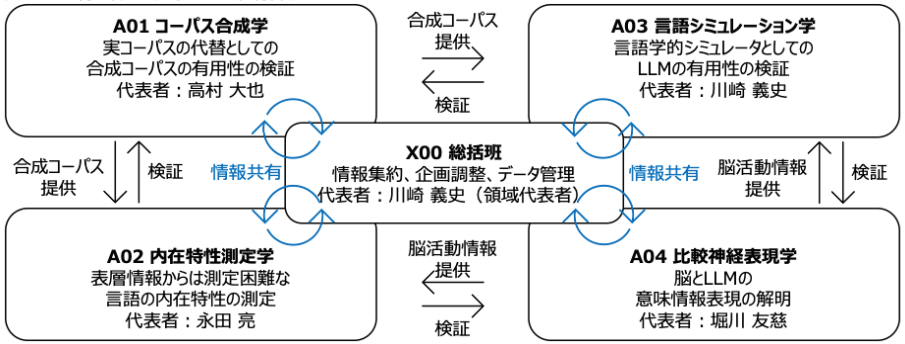

図2　領域構成

ウェブサイト等	領域ホームページ：https://www.insilicoling.c.u-tokyo.ac.jp/ X：https://x.com/insilicolx