

# 大規模言語モデルが学ぶ「言語」を生む 人間の意味生成の理解に向けて



谷口 忠大 立命館大学情報理工学部教授

Tadahiro Taniguchi

2006年京都大学大学院工学研究科博士課程修了。博士(工学)。05年より日本学術振興会特別研究員(DC2)、06年より同(PD)。08年より立命館大学情報理工学部助教、10年より同准教授。15年から16年までImperial College London客員准教授。17年より現職。またパナソニック ホールディングス(株)客員総括主幹技師。専門は人工知能、認知発達ロボティクスなど。本の紹介ゲーム「ピプリオバトル」の発案者でもある。主著に『心を知るための人工知能』(共立出版/2020年)、『記号創発ロボティクス』『イラストで学ぶ人工知能概論』(ともに講談社/2014年)、『ピプリオバトル』(文藝春秋/2013年)など。

## 始まりのコミュニケーション

赤ちゃんが泣き出すと、母親は乳房を寄せて、その唇で吸い付かせる。赤ちゃんは安心したように目を閉じて、喉を鳴らす。母親はその姿を見つめる。

これは私たち人間がこの世に生を受けて、一番初めに行うコミュニケーションの一つだ。赤ちゃんの「泣き声」は、母親にとっては「お腹が減っている」という赤ちゃんの状態や、「お乳が飲みたい」という欲求を表すサインとして発せられる。ここでのサインとは(後に述べるように)記号を構成する要素であり、何かを表す役割を果たすために表出される信号である。

やがて赤ちゃんは乳児から幼児へと成長し、この世界で動き出す。この世界で自律的に生きていく探索者として。クッションの上に「おすわり」するようになった赤ちゃんは机の上のおしゃぶりに手を伸ばす。そして「うーうー」と声を出す。それを見たあなたは「ああ、おしゃぶりを取ってほしいのかなあ」とその意図を読み解く。伸ばした手を意味付ける。伸ばした手は何かを指し示すサインとなる。そうやって「はい、どうぞ」とおしゃぶりを赤ちゃんに渡すのだ。ここで指標記号として表出されたその手をサインとして用いたコミュニケーションは完結し、赤ちゃんのコミュニケーションの目的は達成される。あなたの能動的で創造的な解釈行動とともに。赤ちゃんはそれを柔らかな手で受け取りキャッキヤと喜ぶ。

このようなサインの表出と能動的な意味付けにこそ、コミ

ュニケーションの始まりがある。AI時代になってもそれは変わらない。コミュニケーションとは解釈者の能動的で創造的な意味付けにその基盤を持つ。より複雑な記号としての言語であってもそうなのだ。テキストの文字列そのものに意味があるわけではない。意味とは解釈者が見出すものであり、能動的に与えるものだ。

本稿では、このようなコミュニケーション観を基礎に持って、AI時代のコミュニケーションを考えてみたい。ChatGPTをはじめとした大規模言語モデルに基づく生成AIは人間と機械、人間と人間のコミュニケーションのあり方を大きく変えようとしている。その時代にあって、改めて人間とAIがわかり合うこと、AIが言葉の「意味」を理解することについて、考え



乳児と母親の間の非言語的なコミュニケーション。乳児の泣き声を母親は能動的に解釈してその「記号」の意味を理解する。それは乳児の意図を表しており、母親がそれに対応した行動を取ることで、コミュニケーションは完結する。乳児は泣くことにより母親の行動を引き出している

てみたい。

## 第3次AIブームから生成AIの時代へ

2010年代を通してAI技術は革命的な進化を遂げた。日本国内では第3次AIブームと呼ばれることもある。筆者自身、その渦中においてAIを取り巻く言説の変化を感じてきた。そして時代は2020年代へと至った。AI時代はさらなる変化を遂げようとしている。「認識AI(もしくは識別AI)から生成AIの時代へ」と言っても良いだろう。

画像を入力情報として受け取り、限られたカテゴリ番号やクラスラベルを予測するのが典型的な認識AIだ。画像は3万画素の写真なら、各ピクセルに赤青緑の明るさを表す数字が並び、結果9万次元のベクトルだ。これを例えば3,000個程度の物体クラスへと対応させるのが典型的な「認識」だった。それはいわば9万次元ベクトルという高次元の表現から情報量を切り落とし、人間が事前に設計した3,000次元の離散的で記号的な世界へと落とし込むということをタスクとしている。そのような知的情報処理装置が認識AIだ。これに対して生成AIは言語的な入力を受け取り、その言語が表している画像を生成するといった、より情報量の大きな出力を行う。それは人間が正解を決められる類のものではなく、より(ある意味で)創造的に出力を生み出す。

生成AIと認識AIの間には大きな溝があるわけではなく、技術としては連続的な系譜の中にある。どちらも共にデータを予測することを旨としている。

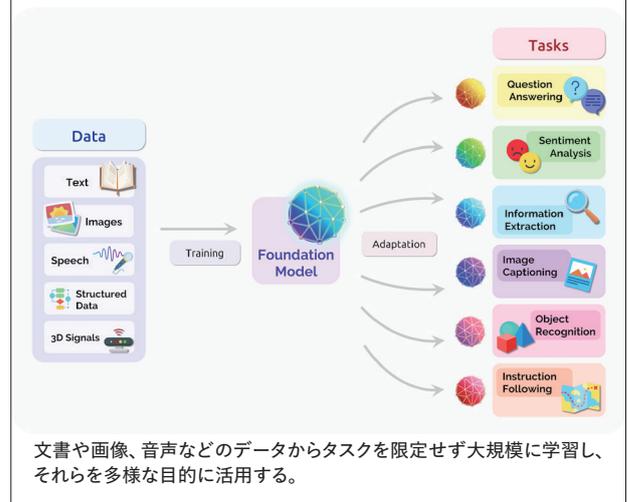
2022年を通して世の中を画像生成AIが賑わせた。2022年の中頃からアニメの女性イラストを生成するようなモデル(NovelAIなど)が出てきてから国内でもSNS上で既存のクリエイターとの軋轢も生まれ始め、社会問題へと発展している。しかし、その中でもこの社会に破壊的かつ創造的なインパクトを与えたのが対話型の生成AIであるChatGPTだろう。

## ChatGPTと基盤モデル

ChatGPTはOpenAIが2022年11月にリリースしたサービスだ。それまでにOpenAIはGPT-1、GPT-2、GPT-3として同社の言語モデルを段階的に進化させており、GPT-3辺りの段階で、AIの研究コミュニティでも多くの研究者が、それまでのAIとの質的な違いとその広がりを感じ始めた。GPT-3をはじめ、それまでになかった大きな規模で言語データを学習した言語モデルは大規模言語モデルと呼ばれる。先に述べた画像生成AIや、その他の画像認識モデルや音声認識モデルと並んで、大規模言語モデルは基盤モデル(Foundation Model)の一種として捉えられる。

基盤モデルに関してBommasani氏らがまとめた「On the

【図表1】基盤モデル(Foundation Model)の概念図



出所: Bommasani et al. "On the Opportunities and Risks of Foundation Models." 2021より引用

Opportunities and Risks of Foundation Models」という論文が話題になった<sup>(1)</sup>。それまでのAIの多くはタスクを決めてそのタスクのためにAIを訓練していたが、基盤モデルは訓練後に解きたい特定のタスクを決めずに、大量のデータを用いてそのAIを訓練する。そして訓練後のモデルをそのまま使って何らかのタスクを行ったり、少量の追加訓練だけでタスクを実行したりする。ある意味で言語についての基盤モデルが大規模言語モデルだ【図表1】。

ChatGPTは大規模言語モデルに基づく対話型AIだ<sup>(2)</sup>。ChatGPTはさまざまな発話に対して、まるで人間のように応答する。その質問応答の範囲は、簡単な会話にとどまらない。むしろツールとして使うことができ、日本語を入力して「英語に翻訳してくれ」と言えばしてくれる。データを放り込んで「表形式にまとめてくれ」と言えばまとめてくれる。「エクセルで〇〇を実行するにはどうしたらいい?」と聞くと高確率で正しい情報を教えてくれる。日英翻訳や英文校正の能力は、それまでに存在していたそれら翻訳・校正の専門AIを凌駕する。ChatGPTの応答の自然さに関して、多くの人々が「人間レベルだ」と驚いてきた。実際にOpenAIのほかいろいろな研究機関の論文が、ChatGPTの性能が標準的な人間の性能を多くのタスクで超えていることを示している。

私たちは人類史上で初めて、人間以上のレベルで言語を扱う非人間的存在(動物や機械)を目の前にしているのである。言語こそが人間と非人間的存在を分かちつものであり、人間の尊厳の根拠の一つであるところまで思いついていた人々にとって、この事実がもたらすインパクトは計り知れない。

ChatGPTの特別さはその参入障壁の低さにある。あらゆる人がブラウザ上でただ言語で問かけるだけで、幅広い用途に用いることができる。この使用の容易さから2022年末のChatGPTのリリース後、その影響は狭いAI分野や情報分野の

中にとどまらずに広がっている。これまでの多くのAI技術と異なり、ChatGPTを使うのに特別な知識はいらないのだ。プログラミング能力もいらない。ただ日本語なり英語で指示文を書けばChatGPTがそれを解釈して、何らかの返答をする。

ChatGPT(もしくは他の生成AI)への指示文はプロンプト(prompt)と呼ばれる。もちろんChatGPTが何を返してくるか(生成してくるか)は入力に依存する。これは人間同士のコミュニケーションにおいて、誰かから情報を引き出そうとするときに、その情報が私たちの質問に依存するのと同じだ。道に迷ったとき、道行く誰かに声をかけ、道を尋ねようにも、どう聞いたら良いかわからない。特に海外においては言語の壁もあり困難だ。海外でなくても初対面の相手に自分の状況と知りたいことを過不足なく伝えるというのはなかなか難しい問題である。ChatGPTにおいても同じである。ChatGPTは大規模な言語モデルの中に知識を持つ。その知識をどう引き出すかは質問にかかっている。つまりプロンプトの選択である。ChatGPTをはじめとした大規模言語モデルや、言語モデルを活用した画像生成AI等に対して所望の振る舞いをしてもらうために、入力するプロンプトを工夫することはプロンプトエンジニアリング(prompt engineering)と呼ばれる。

先にも述べたように、ChatGPTのような大規模言語モデルに基づくシステムでは、ある意味でタスクは事前に定められない。タスクは質問、つまりプロンプトの中に埋め込まれるのだ。「『おはようございます。今日もいい天気ですね』を英語に翻訳してください」のように。タスクの記述さえも、その入力の中に含まれるという、言語の再帰的な性質に気づき、これまでの認識AIにない形で使い、汎用的に広範囲にわたるタスクを解けるようにしたことが、大規模言語モデルがもたらしたブレークスルーの本質の一つである。

余談ではあるが、筆者は従来の認識AI的な言語処理システムの構築方法に本質的な違和感を覚え続けていた。従来のアプローチでは、自然言語処理においてもタスクを定義して、そのタスクの範囲内で自然言語処理装置を個別訓練していた。しかし、私たち人間は言語を学ぶときにタスクなど意識しないし、そもそもタスクのために言語を学習するというようなことはしない。その意味で、ChatGPTのような大規模言語モデルによるアプローチは、私たちがようやくたどり着いた人間の言語に対する「真っ当な」アプローチなのだ。

## ChatGPTは言葉の「意味」を理解しているのだろうか？

確かにAIは言語を扱えるようになった。それを認めるとして、ここで質問を変えよう。

「ChatGPTは言葉の意味を理解しているのだろうか？」

この質問に対しては、さまざまな人が、さまざまな答えを提

示する。思想的もしくは哲学的な立場から否定的な見解を述べる人もいれば、肯定的な見解を述べる人もいるだろう。

典型的なAI研究者であれば、この問いに対してどのような対応をするだろうか？

あくまで筆者の視点でのAI研究者像に基づけば、その研究者は何らかのタスクを探そうとするだろう。「意味理解」ができていなければ解けないと考えられる言語処理タスクを設定して、それで「言葉の意味を理解している」という評価の代替としようとするのだ。

これは意識的にも無意識的にも、AI研究が抱える機能主義と、知能情報処理として知能を見る視点を提示している。これはある意味で、AI研究の教義(ドグマ)のようなものである。前者の機能主義とはいわば「知能とは何かができるということ」として見るし、測るという態度であり、後者は「知能とは入力に対して出力を返す情報処理である」と見る態度である。その範囲において「言葉の意味」を議論すると、言語を入出力とした何らかのタスクが解けるかどうかで見るしかない。そこには意味を感じる主体という視点、意味を生み出す主体としての視点が欠落している。その機能主義的で情報处理的なドグマにより、主体による「意味」生成は欠落せざるを得ないのだ。では、「言葉の意味」とはなんだろうか？

言語は記号の一種である。記号の意味作用に関して議論するパースの記号論では、記号をサイン・対象・解釈項の三項関係と見なす。一般的な自然言語を例にとって説明すると、サインは文字列や音列としての言葉そのものであり、対象とはそれが指し示す物事である。これに対して、解釈項はその2つをつなげる第三項であり、例えば解釈者の意識内でのイメージである。

この三項関係が生じたときに、サインは記号になり、そこに意味が生まれる。言語のほかにも記号論ではさまざまなものが記号になりえる。ベッドの上に置かれたぬいぐるみが天折した娘のことを思い出させれば、そのぬいぐるみは娘の不在を意味するサイン(記号)となるだろう。戦国時代において、山の向こうに見える狼煙は、敵軍が攻め込んできたことを意味するサイン(記号)となるだろう。

重要なのは、サインはそれ自体では意味を持たず、常に解釈主体の能動的な意味付けにより意味を持つということである。

ここでパースが導入した初歩的な記号の分類学を紹介したい。パースは記号の分類学を発展させて、さまざまな三分法を導入したが、サインと対象の関係に関する三分法が最も有名であろう。それは、類像記号(icon、アイコン)、指標記号(index、インデックス)、象徴記号(symbol、シンボル)の3つである。これらはサインと対象がどのような力によって結び付けられているか、ということに関する三分法である。

類像記号ではサインの質そのものが対象と関係付く力を持っている。例えば写真はその見た目によって、その被写体のサ

インとなる。指標記号は物理的因果関係などの必然的な関係性によってサインと対象を結び付ける。足跡が誰かの存在を意味するサインとなることや、指差しがその指し示す先を表すことなどである。一方で、象徴記号は文化や規約によって対象と結び付く。言語は基本的に象徴記号である。例えば「りんご」が赤い果物のりんごを表す記号であることに必然性はない。赤い果物を叩いても「りんご」などという音は鳴らないし、物理的な因果関係も「りんご」という音とその物体の間に存在しない。その証拠に英語では同じ物体に「Apple」という全く異なるサインが割り当てられる。このような任意性は記号の恣意性と呼ばれる。

ここで重要なのは、記号の意味とは解釈者の中で生まれるものだという点である。もちろん、その言葉が何を表すかということに対する社会的了解は存在している。その言語圏(より一般的には記号圏、semiosphere)において生きていく中で、主体はその解釈をその言語圏の解釈に合わせていこうとする。ただしその言葉の意味は、社会から一方的に私たちに押し付けられるものではなく、私たちの「意味付け」に関する不断の努力の上に成り立っていることを理解せねばならない。言い換えれば、記号に意味を与え、社会において言葉の意味を維持することに私たちは多大な費用(コスト)を支払っている。母親が赤ちゃんの泣き声や幼児の指差しを解釈するような能動的な「意味付け」という行為によって。「意味付け」は無料(ただ)ではなく、私たちの無意識な努力の積み重ねにより、言葉の意味は維持されているのだ。

ここで冒頭で述べた赤ちゃんの例に立ち返ろう。赤ちゃんの「泣き声」はサインだ。母親はなんとかそこに意味を見出そうとする。赤ちゃんが求めているだろうことを想像しながら、乳房をその口元にそっと寄せるのである。赤ちゃんが手を伸ばしているときだって同じだ。その動作に何らかの意図を見出し、おしゃぶりを取ってあげるのである。手を伸ばすことは一般的な「指差し」と同様に何かを指し示す指標記号の一種であり、この一連の意味付けが指標記号の立ち現れに対応している。やがて指差しという指標記号は、手話という象徴記号の言語的体系の一部に組み込まれていく。

お互いに他者の動きを補完し、意図を汲み取り、それぞれが見てきたことの情報を共有し、協調することで、この世界で生きていく。記号は、私たちがこの世界で生きていくために生み出していくものだ。意味ある記号系とは、自律的な存在である生物としての私たちが、不断の意味付けの努力を統合することで、社会の中で創発するものなのである。

## ネオ・サイバネティクスと 記号創発システム

このような「自律的な存在である生物としての私たち」を基

礎に置いて、言語を捉えていくためには、私たちはテキストとしての言語ありきの世界観、そして知能を知的情報処理としてみる世界観から脱さねばならない。むしろ「世界観」より「システム観」という言葉を用いるほうが妥当かもしれない。

ネオ・サイバネティクスの系譜にある基礎情報学を展開する西垣通氏は情報概念に対して、三分法を導入した<sup>(3)</sup>。機械情報と社会情報と生命情報である。詳細には立ち入らないが、簡単に説明する。

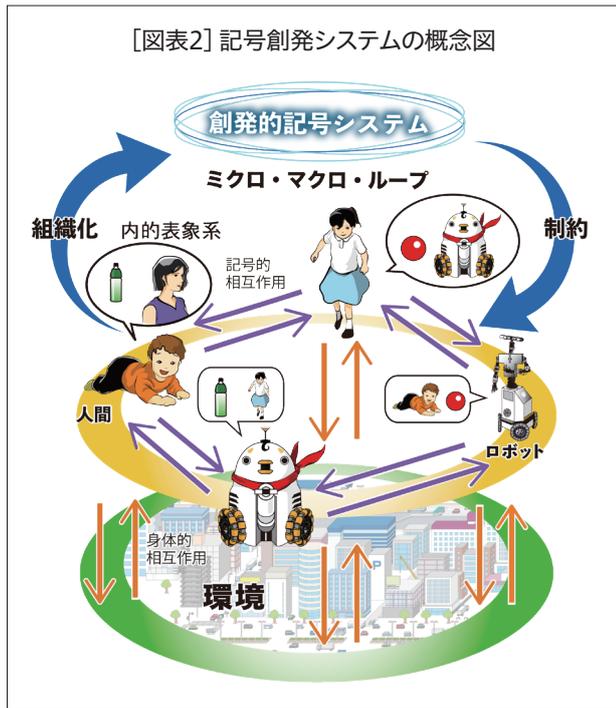
機械情報とはいわゆる数理的な情報科学で議論することの多い情報であり、典型的にはコンピュータの中のビット列として表される情報である。そのビット一つ一つに何らかの「意味」が担わされていることは、必ずしも必要ない。それらが状態として区別され、ルールに基づいて計算処理されることが重要である。社会情報は書籍に書かれた文書のようなものを考えればよい。本稿も社会情報の一種であろう。もちろんある意味での「意味」を文章は持っている。しかし、それが真に意味を持つのは、各読者がそれを読み理解したときであろう。生命情報は自律的な生命としての主体にとっての意味である。

ネオ・サイバネティクスの議論の中で上述の西垣氏や原島大輔氏は、情報処理的なシステム観(コンピューティング・パラダイム)からサイバネティックなシステム観(サイバネティック・パラダイム)への移行の必要性を主張する<sup>(4)</sup>。知能の話に寄せて考えよう。サイバネティック・パラダイムはシステムを局所的な主観性から観察する。つまり主体にとって現れる世界を問題にするのだ。この世界に生きていく人間の「知能」とは、決して入出力変換で表される情報処理ではなく、自律的に存在し続けるシステムであると考えているのだ。コミュニケーションの意味を問うためには、私たちはただ言語の情報処理を見るのではなく、主体による記号の生成と、その能動的な解釈、そしてそれに基づく環境適応を考えていく必要がある。

記号創発システムは、このような記号観を統合する「意味」生成のシステム描像である<sup>(5)</sup>。各主体は他者と協調しながらも身体的相互作用を通して、環境に適応しよう、生き抜こうとする。その中で世界を認識するための内部表現(もしくは内的表象)を得ていく。またそれらの内部表現に基づいて、サインを生成し、それを他者が意味付けすることで、記号的相互作用(記号や言語を用いたコミュニケーション)が生まれていく。記号が意味を持つためには、主体による意味付けのみならず、言語がいかにかに用いられ、何を表すのかについての規範に関する合意が社会の中で必要となる。それが創発的な記号体系として、その社会の中で立ち現れてくるのだ。それが個々の主体に学習され、またその発話と解釈に制約を与えることで、記号(特に象徴記号)は社会の中で意味を持つようになっていく[図表2]。この全体のプロセスはどのようなメカニズムであり、それはどのような数理モデルとして表現可能なのだろうか？

筆者らは人工知能分野における機械学習の数理モデル、特

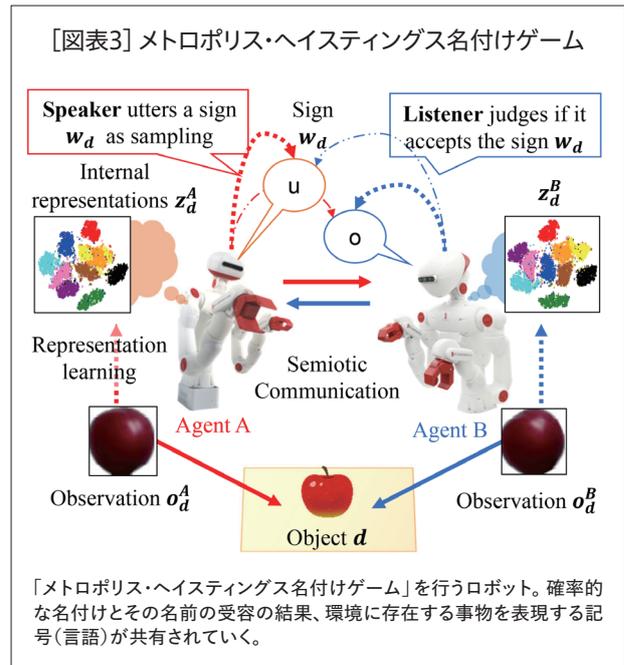
[図表2] 記号創発システムの概念図



に確率的生成モデルによって定式化される表現学習に基づいて、記号創発システムの構成論を提案している<sup>(6)</sup>。ここでは、各エージェントが環境の知覚情報(簡易な例としては視覚情報)を予測できるように学習をしていく。人間の認知を「予測」を軸に説明することは予測符号化(predictive coding)や自由エネルギー原理(free energy principle)という名称で、近年、認知(もしくは脳)の一般理論として注目を集めている。しかし、それらは基本的に各エージェントを主体とした表現学習である。ここで大胆に、表現学習(もしくは予測学習)の主体を、各エージェントからそれらが共有する「言語」に移す。その「言語」が各エージェントの内部表現を介して、各エージェントの知覚情報を予測すると考えるのだ。もしそれが可能なのであれば、「言語」は外的表象として全エージェントの知覚情報を統合するような役割を果たす。果たしてそのようなことが現実に起こりうるのだろうか？

筆者らは「メトロポリス・ヘイスティングス名付けゲーム」という、ある種の言語ゲームを提案した<sup>(6)</sup>。各エージェントが自らの内部表現に基づきサインを生成し、またその受け手が自らの内部表現に基づき計算される、ある種の確率に基づいて自律的にそのサインを信じるか信じないかを判断する。実は一定の条件が満たされれば、このような名付けゲームを通して、エージェント同士が記号を共有し、さらに記号の創発を通して集団の観測を統合し、表現学習が行われることが示された[図表3]。つまりこれは全エージェントの認知の統合が実現されることを意味する。この結果に基づいて、筆者は、社会における記号創発が集团的予測符号化として捉えられるのではないか、という仮説を提出している。この考え方はまだ仮説の段階ではあるが、記号創発システムのダイナミクスを探索

[図表3] メトロポリス・ヘイスティングス名付けゲーム



「メトロポリス・ヘイスティングス名付けゲーム」を行うロボット。確率的な名付けとその名前の受容の結果、環境に存在する事物を表現する記号(言語)が共有されていく。

することが、生命情報を基盤にした言語の本質理解——そして意味の理解につながると考えている<sup>(7,8,9)</sup>。

## まだまだ子どもな大規模言語モデル

ChatGPTのリリース以来、大規模言語モデルの使用が一般に広がって、半年以上が経った。爆発的な利用拡大は、人間社会の言語活動にも大きな影響を与えている。特に教育業界において危機感が強い。授業のレポートやプログラミング演習の課題を学生が生成AIを用いて作成した際に成績評価をどうするのか？ 禁止するにしても、そもそもそのレポートやプログラムが生成AIで作られたものと判別することができるのか？ 教育分野にとどまらず、既にインターネット上には、それと知らない間にChatGPTで書かれた文書が人間により書かれた文章のふりをしてアップされている。ここで新たな問題が生じ始める。AIカニバリズムだ。AIカニバリズムとは、生成AIが生成した文書をまたAIが学習データに用いてしまうことで、どんどんAIの能力が低下していくことを意味する。

全世界の人間たちが、それぞれの意味付けの努力に基づいて書き続けてきた文書。それを大量のコーパスとして学習することで、大規模言語モデルは言語処理の機能を得ている。いわばChatGPTは人間の生成したテキストという栄養を摂取することによって、その能力を得ているのだ。一般的にChatGPTは、この世界で自律的に生きていくための「身体」は持たない。ChatGPTが内化する言語は、人間が環境適応の中で、また他者との協調のために生み出してきたものだ。人々が記号創発システムの構成要素となり、その感覚運動情報を予測符号化し内部表現を得て、分節化することで記号と対応付け、それに基づき他者とともに集团的予測符号化を通して、社

会の中で言語を形成していくのだ。コーパスの中にある発話(や生成された文)はそのような言語からのサンプリングの結果であり、その全ては冒頭で述べたような意味付けの主体としての人間の言語活動に支えられている。記号創発システムの図に示すように、言語とは創発的な存在ではあるが、その下層には人々の言語活動があり、その上澄みとして私たちがインターネット上に出力するテキストデータがあるのだ。

生成AI——大規模言語モデルの進歩は目覚ましく、上澄みとして蓄積された大規模なコーパスをただ予測学習するだけで、さまざまな言語処理タスクが実現できるようになることが示された。しかし、現在の彼らは彼ら自身が「生きていくための言語」を生み出しているわけではない。大規模言語モデルによる知能は、私たち人間の意味付けする力、記号を生み出す力によって生まれた言語に依存している。つまり私たちの記号創発の力に依存している。私たちが世界を分節化し、言語活動を行えるからこそ、そしてその結果生まれた大規模なコーパスを提供するからこそ、大規模言語モデルは機能を持てるのである。AIカニバリズムは、生成AIが現状のままではまだ記号創発システムを支え続ける構成員にはなれないことを示唆している。

社会で働き価値を生み出す大人と、大人が生み出した財やサービスをただ消費する子どもの関係が、人間と大規模言語モデルの比喻として、なんとなく思い浮かべられる。大規模言語モデルは人間が生み出した言語資源をただ受容し消費し、その上で知ったかぶりをしているような雰囲気がある。そういう構造にある。その意味で、生成AIはまだ、広義での人間の言語能力の次元には届いていない。機能主義的な視点では人間に届きつつあるAIも、構成主義的な視点では人間に届いていないのだ。

## これから

いかに大規模言語モデルの使用がAIカニバリズムを招こうとも、それでも人々は大規模言語モデルの使用を止めないだろう。そして、大規模言語モデルを利用することが当たり前となった人々の活動が、AI時代のコミュニケーションを形作っていく。このような時代においては、なすべき学術研究は少なくとも2つあるように思う。

1つは大規模言語モデルを用いて、人々がコミュニケーションを続けることによる言語と社会の変質を予測し、理解し、それに対応する適切な制度設計や、順応していく私たち自身のあり方を考えていくことだろう。

もう1つは、人間とともに言語(記号)を生み出す活動に参画することができるようにすることだろう。そのためにはAI自体が自律的な主体として、私たちと共にこの世界に参画し、協調のために事物を表現する記号を生成し、またその記号を意

味付ける活動を行っていきけるようになる必要がある。つまり記号創発システムの構成要素となる必要がある。そのためには、機能主義的なコンピューティング・パラダイムを脱し、構成主義的なサイバネティック・パラダイムへと移行することが必要だろう。大規模言語モデルの次のステップへ向かうためには、大規模言語モデルがただ「喰う」だけに終わっている言語そのものを、その意味付けも含めて生成していくことができるようになる必要があるだろう。これは本質的な意味でAIカニバリズムを脱する道ともなる。それは私たちと共に「意味付け」の活動に参画し、記号創発を通して私たちと共にこの世界で生きていくようなAIへの道だ。

人間の知能の本質をどう捉えるのか？ 言語的(記号的)コミュニケーションの「意味」という次元は、私たちに大規模言語モデルが到達する知能のさらにその先に関する思索や洞察を求める。表層的な言語処理や、機能的な意味でさまざまなタスクを生成AIができるようになった現在だからこそ、私たちは人間のコミュニケーションに関するさらなる深層へと踏み込めるのだ。AI時代のコミュニケーションはそれ自体を変容させていくとともに、私たち自身のコミュニケーションの本質をより露わにしていくのだろう。その先に、本当の意味での「AI時代のコミュニケーション」があるのかもしれない。

<註>

- (1) Bommasani, Rishi, et al. "On the opportunities and risks of foundation models." arXiv preprint arXiv:2108.07258 (2021).
- (2) ChatGPT by OpenAI <https://chat.openai.com/>
- (3) 西垣通『基礎情報学—生命から社会へ』NTT出版、2004
- (4) 西垣通『新 基礎情報学—機械をこえる生命』NTT出版、2021。谷口忠大、河島茂生、井上明人(編著)『未来社会と「意味」の境界：記号創発システム論/ネオ・サイバネティクス/プラグマティズム』勁草書房、2023
- (5) 谷口忠大『心を知るための人工知能：認知科学としての記号創発ロボティクス(越境する認知科学)』共立出版、2020
- (6) Taniguchi, Tadahiro, et al. "Emergent communication through Metropolis-Hastings naming game with deep generative models." arXiv preprint arXiv:2205.12392 (2022).
- (7) 谷口忠大「分散的ベイズ推論としてのマルチエージェント記号創発」日本ロボット学会誌、2022、40巻10号、p.883-888、<https://doi.org/10.7210/jrsj.40.883>
- (8) Taniguchi, Tadahiro "集合的予測符号化に基づく言語と認知のダイナミクス～記号創発ロボティクスの新展開に向けて～" PsyArXiv, 8 June 2023, Web.
- (9) "NLP2023 招待講演1「社会における分散的ベイズ推論としての記号創発～集合的予測符号化としての言語観～」谷口忠大先生(立命館大学) 2023年3月15日 沖縄コンベンションセンター" YouTube, uploaded by 言語処理学会 ANLP YouTube Channel, 24 Mar. 2023, <https://www.youtube.com/watch?v=RrrkwBwnbOs>