

意味研究に突破口を開く

認知言語学者が AI 研究に期待するもの

町田 章

はじめに

認知主体の主観的な認識の在り方(=概念化)を言語表現の意味として分析に組み込んだ認知言語学は、ときに、反証可能性がないなどと批判されることがある。そのため、近年では統計学や心理学的実験などの手法を積極的に取り入れることによって分析の客観性や科学性を担保しようとする試みも行われてきた。しかしながら、概念化に重要な役割を果たす認知主体の事態の捉え方(construal)の研究に関しては、十分に実証可能な方法論はまだ確立されたとは言えない。実際、話者が事態をどのように捉えているかに関しては、言語現象を頼りに分析者が推論(推理)するという方法が、一般に、いまだにとられているのである。本研究では、AIを用いた逆心理学(reverse psychology)の手法が、事態に対する話し手の捉え方を含んだ意味に関する仮説を検証する手段となりえ、より実証的な意味に関する研究の可能性が開かれることを示す。

認知言語学と機械学習(深層学習)の親和性

認知言語学のパラダイムを端的に述べるのはほぼ不可能であるため詳細は入門書等に譲るとして、ここであえて二つのポイントを挙げるとしたら、認知言語学では、言語現象をできる限りカテゴリー化の問題に還元して捉えるという点(cf. Taylor 2003)とそのようなカテゴリー化の現象は人間の認知能力と言語使用の観点から説明されるという点(cf. Langacker 1999)が挙げられるだろう。そのため、認知言語学では、いわゆる主語や動詞などの文法項目や構文などの統語現象も、従来の言語学とは異なり、それぞれ認知的なカテゴリーとして扱われ、その成立の動機づけを人間の認知能力と実際の言語使用の中に見出すというアプローチをとっている。これは、言語現象をある種の記号演算とみなし、有効な演算の条件や制約などを探求する従来の理論言語学のアプローチとは一線を画すところである。

一方、人工知能(以下、AI)による自然言語処理研究は、独自の発展を遂げる中で、世界に関する一般的な知識、文脈理解、他者の意図理解の重要性など、言語学が直面してきた問題とほぼ同じ問題にチャレンジしてきた。そのような中で特筆すべきは、AI研究が人間の手によって条件や制約などをプログラミングする時代から、対象に関する膨大なデータの中からAIが自ら対象の特徴を抽出し対象の概念を習得する機械学習の時代に突入したということである。特に、2012年ごろから注目を集めるようになった深層学習(Deep Learning)と呼ばれる機械学習モデルは、画像認識などの分野において革命とも言えるほどの飛躍をもたらしている(cf. 松尾 2015)。人間の脳内の神経回路を模したこの新しいAIの最大の功績は、膨大な事例の学習によってAIが自ら「概念」を学習することが可能であることを原理的に示したことであり、それは大局的な見方をすれば、膨大な具体事例からあらゆる言語カテゴリーを習得するという認知言語学の用法基盤アプローチ(usage-based approach)に強力な支持基盤を与えたことになる。これは、Lakoff and Johnson (1999)が求める収束的証拠(convergent evidence)の一つとして認知言語学にとって重要な意味を持つ。

意味に関する仮説の検証と循環論

仮に言語表現の意味が外的世界の客観的現実(objective reality)を心内に写し取った心内表象(mental representation)であるとしたら、ある表現の意味とされるもの(=仮説)が外的世界の対象と一致していれば仮説が検証されたことになるため、仮説の検証は比較的簡単で客観性を持たせやすいだろう。しかしながら、言語表現の意味には、概念内容(conceptual content)だけでなく、話し手の捉え方(construal)も含まれるという立場をとる認知言語学では、外界世界の客観的現実との一致関係の有無を調べたところで、意味に関する仮説を検証したことにはならない。例えば、(1)の表現の表す意味の差異に関しては、外界にある実物(実際の部屋)は何も語らない。当然のことだが、「丸い部屋」と「丸くなっている部屋」の意味的な差異を生み出しているのは外界にある何かではなく、話し手の心(=捉え方)だからである。

- (1) a. その部屋は丸い。 b. その部屋は丸くなっている。 (Matsumoto 1996:124)

このように、話し手の捉え方までをも意味に含むと考えると、外的世界との対応関係では意味に関する仮説を検証することができないことになる。それ以上に問題なのは、このような考えでは、表現を手掛かりに話し

手の捉え方を推測するため、議論の循環が避けられないことである。例えば、(1)では心的走査(mental scanning)による仮想的な変化(fictive change)を話し手が認識したかどうかの意味の差異を生んでいるという仮説を立てることができるが、なぜそのような仮説が立てられたのかというと、それは、「丸い」部屋に対して「丸くなっている」という動的な表現が用いられたからだということになる。つまり、話し手の捉え方と言語表現は不可分の関係にあるため、捉え方を知るためには表現が必要であり、表現を見ると捉え方がわかるという関係になっているのである。このため、表現から切り離して捉え方だけを検証するのは非常に難しくなっている。

実際、(2)の差異に関して、(2a)では、鉛筆を道具として捉えているのに対し、(2b)では、鉛筆を抽象的な範囲として捉えているという仮説を立てることができると思われるが、それではなぜそのような仮説が立てられるかかというと、(2a)では、道具を表す *with* が用いられ、鉛筆が可算(*a pencil*)として表現されているからであり、(2b)では、範囲を表す *in* が用いられ、鉛筆が不可算の抽象物(*pencil*)として表現されているからであるということになる。ここに表現と捉え方の循環論が見て取れる。

- (2) a. I wrote the letter with a pencil. b. I wrote the letter in pencil.

AI による逆心理学の可能性

以上のような議論の循環が避けられないのは、つまるところ、我々は心内表象(mental representation)を直接観察することができないという問題につく。表現の意味が表す心内表象は客観的現実ではなく、心的現実(psychological reality)を表したものであるため、この心的現実を何らかの方法で検証しなければならないのである。そして、この困難な状況に対して僅かながら希望の光をもたらしてくれるのが、機械学習によるシミュレーションなのである。結局、意味の検証が難しいのは、意味が心的現実を表しているからであるが、これは、逆に言うと、心的現実が AI によってシミュレーションできれば意味の検証が可能になるということになる。

実際、AI によるシミュレーションを用いて人間の心理現象を明らかにする逆心理学(reverse psychology)という研究手法が誕生しつつある(Watanabe et al. 2018)。Watanabe et al. (2018)の深層学習による錯視画像の生成実験では、機械学習によって学んだ AI は人間と同じように錯視を起こすことが示されている。錯視とは、実際には動いていないものが動いているかのように知覚されるなどの現象であるが、これは単なる人間の思い違いとして片づけられない。たとえ動いて異なることを十分に知らされていたとしても、それでも、人間の脳はその対象が動いているかのように知覚処理してしまうからである。言い換えると、錯視は、客観的現実と心内で表象される心的現実が何らかの理由で乖離してしまった現象なのである。そして、この実験が示唆していることは、AI が機械学習によって獲得する表象は客観的現実ではなく心的現実と一致しているということである。AI が心的現実を表象しているのであれば、テキストからの画像生成(text-to-image generation)と呼ばれる技術で生成される画像はテキストが表象している心的現実(の一部)である可能性が高い。そうであるならば、具体的な表現をもとに生成された画像は、話し手の捉え方までも含んだ意味(概念化 conceptualization)の一部ということになる。つまり、テキストからの画像生成技術を用いれば、近い将来、「丸くなっている部屋」が表象している仮想変化をシミュレーションし、直接検証可能な画像として提示してくれるかもしれないのである。

まとめ

言語表現の意味を概念内容とその捉え方からなると考えた場合、表現の意味が心的に表象しているものは必然的に心的現実であるということになる。そのため、心的現実を直接検証する手段が求められるが、これまで、心理学実験などを用いて部分的に間接的に検証するしか手段がなかった。それに対し、今後の AI の発展次第では、AI を用いた逆心理学(reverse psychology)の手法を用いて心的現実をシミュレーションすることが可能になっていくかもしれない。この手法が発展すれば、主観主義的意味論が抱える表現と捉え方の循環論を脱却することができ、意味というブラックボックスの中身を観察可能にすることができるかもしれない。

参考文献

Lakoff, G. and M. Johnson (1999) *Philosophy in the Flesh*, Basic Books. / Langacker, R. (1999) *Grammar and Conceptualization*, Mouton de Gruyter. / Matsumoto, Y. (1996) Subjective-Change Expressions in Japanese and Their Cognitive and Linguistic Bases, in Fauconnier, G. and E. Sweetser (eds.), *Spaces, World and Grammar*, The University of Chicago Press, 124-156. / Taylor, J. R. (2003) *Linguistic Categorization* 3rd ed, Oxford University Press. / Watanabe, E., A. Kitaoka, K. Sakamoto, M. Yasugi and K. Tanaka (2018) Illusory Motion Reproduced by Deep Neural Networks Trained for Prediction, *Frontiers in Psychology* 15, 1-12. / 松尾豊 (2015) 『人工知能は人間を超えるか』 Kadokawa.