

## 理論言語学に未来はあるか？

窪田悠介（国立国語研究所）

要旨：深層学習の急速な発展などにより、広い意味での言語理解の問題への関心が AI 研究の分野で再び高まっている。機械翻訳の飛躍的な精度向上に驚かされたのはつかの間、今や深層学習の手法の理論言語学の問題への応用の試みなども出始めている。我々言語学者はこの状況にどう対処すべきなのか？ 本発表では、まず AI 研究と対峙する理論言語学研究の現状を概観する。ここから、理論言語学の本来の課題であるはずの「習得可能性」の問題にまつわる研究史的問題が明らかになり、それを振り返ることで、理論言語学が現在抱える大きな課題が浮き彫りになる。そして、この課題に真剣に取り組むためには認知科学の中に AI 研究の進展を有機的に位置付ける大きな視点が必要となることもほぼ必然の帰結として導かれると思われる。理論言語学に未来があるかどうかは、言語学者がこの事実を虚心坦懐に直視することができるかどうかにかかっている。

理論言語学にかつてほどの活気が感じられない。自然言語処理の分野においては、深層学習の発展により構文解析や機械翻訳はすでに「終わった」課題とされ、工学的な応用におけるシステムの精度向上に言語学の知見が役立つ可能性は完全に閉ざされてしまった。深層学習というのは、かつて理論言語学者が完膚なきまでに叩き潰したはずのコネクションスト研究の末裔である。言語学の世界の外側では、いつの間にか経験主義的な立場が席卷し、合理主義的な立場は影を潜めている。「経験主義の逆襲」の勢いは言語学の外の世界にとどまらない。現代的な深層学習の手法で理論言語学の核心的な問題にどこまで迫ることができるかを問う研究も既に現れはじめている。我々は今や最後の砦さえ失おうとしているのか？

言語学者にとって、最後の砦とは何だろうか？ それは言うまでもなく、チョムスキーが生成文法の課題として掲げた「習得可能性 (learnability)」の問題である。たとえ機械翻訳の精度がどれほど完璧に近づいても、そして、たとえ冗談が分かるほどに言語の意味をよく理解するロボットが発明されたとしても、人間の子供が母語を獲得するメカニズムが解明されなければ、理論言語学が自らに課した課題が解かれたということにはならない。これは、いわゆる「統語論の自律性」を標榜する狭い意味での生成統語論研究者だけでなく、意味(や音)の問題に取り組む理論研究者も広く共有する見解であろう。ここで「解明」というのは、単にそのシステムを模倣する(中身がブラックボックスの)モデルを作るのではなく、人間の研究者に(少なくとも原理の骨格部分は)理解可能な形で母語獲得の過程に関わる根本原理がつまびらかにされる、という意味である。これは、より実用に近い工学的な AI 研究の中心課題ではないかもしれないが、認知科学の究極的な課題の一つであることは間違いなく、その意味で、「知能」を人工的に模倣することを目的とする AI 研究にとっても、隣接分野における最重要の問題の一つと位置づけられる。ひょっとしたら、そもそもこの課題は、上で述べたような意味での「解明」が不可能なのかもしれない。(言うまでもな

く、言語の社会的、文化的、歴史的側面を潔く捨象するという大胆な単純化により、言語習得の問題を明確に定義した点にチョムスキーの天才たるゆえんがあったわけだが、そもそも問題設定に無理があったのかもしれないのだから。) だが、もし不可能なのだとしたら、何故不可能なのかに対して合理的な説明が与えられなければならないし、「習得可能性」の問題を他の形で理解する方法を我々は見出さなければならない。子供が明示的な教示なしにノイズが多く不完全なデータ（と、とりあえずは考えてよいであろうもの）から母語を短期間のうちに獲得するという事実は、まぎれもなく事実として存在するのである。

とはいえ、理論言語学自体も、必ずしも習得可能性の問題自体に正面から取り組んできたとは言い難い。逆説的ではあるが、習得可能性のテーゼは、（その解明に向けてどのように問題を切り分けていったらよいかという問題意識が分野全体で共有され、真剣に議論される）具体的な解くべき課題として位置づけられるというよりは、生成文法理論（や、それと共に発展してきた自然言語の意味論の理論）が極度に技術的になっていく過程で、抽象的な理論的構築物を正当化するための盾としてしばしば用いられてきた。この典型的な例を、統語論研究における「島の制約（island constraint）」の位置づけに見ることができる。「島の制約」というのは、簡単に言うと、文の純粋に構造的な側面に関する極めて抽象度の高い制約によってある種の *wh* 疑問文などが容認不可能となる現象であり、生成文法研究においては、このような制約が自然言語に関する普遍的な制約として存在することが、文法能力が生得的な知識であることの強い証拠であるという説が定説となっている。しかしながら、単にこのような知識が生得的だと言っただけでは、母語の獲得の具体的なプロセスは何一つ解明されないままである。理論言語学研究においては、この、中心的な課題であるはずの問題に対して、過去 60 年間特に大きな進歩はなかったと言ってよい。

さて、理論言語学に未来はあるか？ 今までのように習得可能性の問題を都合のよい隠れ蓑にし、問題を棚上げにし続ける限り、理論言語学にはもはや未来はないように思われる。この場合、我々は最後の砦を（形だけは）守ることが出来るだろうが、「習得可能性」の問題はその重要性が理解されないまま打ち捨てられ、（少なくとも当分の間は）そもそも解明が試みられることすらなくなってしまいうだろう。だが、未来はそれほど暗くないと筆者は考えている。（習得可能性の問題が、単なるドグマとして歪曲・矮小化されて通用してきた側面が少なからずあるとはいえ、）過去 60 年の理論言語学研究の成果として、少なくとも、「何が解明されなければ人間の言語習得・言語理解の問題に答えが出たとは言えないか」ということに関しては、ある程度問題の整理が進んだと考えてよいように思われる。上で例に出した「島の制約」は、まさしくそのような問題の整理の一つと位置づけることができる。このような観点から理論言語学の未来を考えたとき、どのような可能性が見えてくるだろうか？ そして、理論言語学は急速に進展する AI 研究とどのような関係を持つことができるだろうか？ これらはいずれも、理論言語学が現在分野として直面している大きな問題である。本発表では、このような問いを考え始めるためのきっかけを多少なりとも提供することを試みたい。