

令和3年8月19日

報道機関 各位

東北大学大学院情報科学研究科

## 言葉を処理するAIは人のように「読み滞る」か 高性能なAIと人の認知の間に乖離を確認

### 【発表のポイント】

- AIの言語処理と人間の読み活動（読み時間など）が似た傾向を示すかという観点から、AIの認知的な妥当性が分析されてきた。
- AIの言語処理の認知的な妥当性については、英語を対象とした研究がほとんどであり、言語普遍的な知見が得られているのか明らかでなかった。
- 言語処理モデルの認知的妥当性とその工学的な性能の間に単調な関係が成り立つという報告が、英語とは異なる言語で必ずしも成り立たないことを示し、言語普遍的な知見が得られていないことを指摘した。
- 言語間での結果の乖離について一つの理論的な解釈を与えた。

### 【概要】

人間が言語を理解する際にどのような処理が行われているか、その計算メカニズムは未だ謎に包まれています。近年、言葉を処理するAIの技術を活用し、人間の言語処理について洞察を得る試みがなされてきました（図1）。しかしながら、そのような研究のほとんどが英語を対象にしており、これまで得られてきた知見の言語普遍的な一般性は不明でした。

東北大学大学院情報科学研究科の栗林樹生研究員（博士後期課程、日本学術振興会特別研究員）、乾健太郎教授、東京大学の大関洋平講師、国立国語研究所の浅原正幸教授らの研究グループは、AIの認知的妥当性について日本語と英語を対象とした実験を行い、既存研究で報告されていた知見が言語普遍的な一般性を欠く可能性を指摘しました。特に近年、言語を扱うAIの工学的性能は飛躍的に向上し続けており、認知科学的な視点から、この発展が何を意味するのかという学際的な問いに対して示唆を与えている点で意義のある成果と言えます。本研究結果は、自然言語処理分野内で最もインパクトの大きい査読付き国際会議<sup>1</sup>である Annual Meeting of the Association for Computational Linguistics (ACL2021) に採択され、8月4日に発表されました。

<sup>1</sup> [https://scholar.google.co.jp/citations?view\\_op=top\\_venues&hl=ja&vq=eng\\_computationallinguistics](https://scholar.google.co.jp/citations?view_op=top_venues&hl=ja&vq=eng_computationallinguistics)

### 【詳細な説明】

自然言語処理分野では、日本語のような人間が扱う言語を処理できる AI について研究されており、翻訳など一部の問題設定では人に匹敵する性能を達成し始めています。一方、認知科学・計算心理言語学の分野では人間がどのように言語を処理するのか、人間の言語処理について研究が行われています。本研究では、自然言語処理で開発された AI（言語モデル）について、認知科学・計算心理言語学の視点からその認知的妥当性を調べ、同時に人間らしい言語モデルを実現するための要件を整理することで、人間の文処理の計算モデルについて洞察を得ることを目的としています。具体的に本研究では、文を読んだときの人間の読み活動データ（視線移動）と言語モデルが示す統計量（サプライザル）を比較することで、言語モデルの認知的妥当性を測定しています（図1）。このような人間の読み活動データと言語モデルの挙動を直接比較する研究のほとんどは英語を対象に行われており、これまでの研究からどれほど言語普遍的な知見が得られているのか定かではありませんでした。

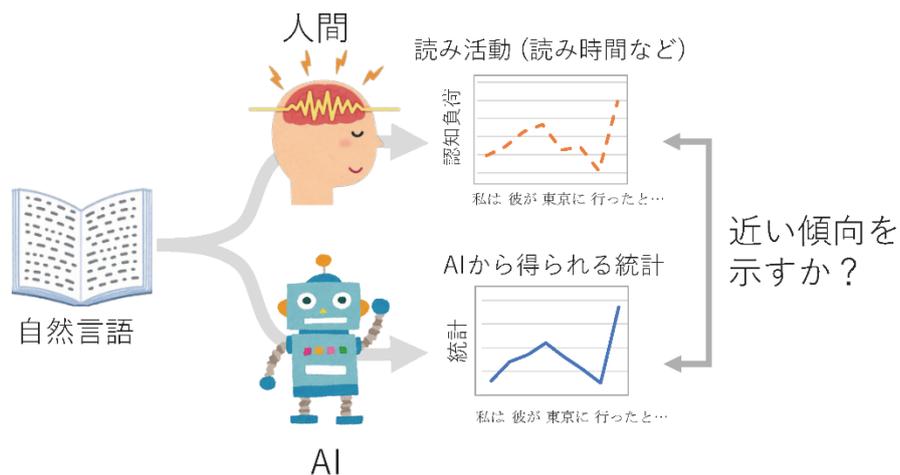


図 1: 人間の読み活動と AI のふるまいの比較

本研究では、英語と日本語を用いた対照実験を行い、既存の知見の言語横断的な一般性に対して示唆を与えることを目的としました。特に「工学的な性能が高いほど言語モデルの認知的妥当性も高い」という既存の定説に反する結果が得られ（図2）、言語間で観察が異なる一つの理論的な解釈も与えました。特に近年、言語モデルの工学的性能はモデルを巨大にしていくことで向上し続けるという「スケーリング則」が提示されており、モデルを巨大にし続けた先に人の認知モデルが存在するののかという分野間の未来に対して示唆を与えている点で意義があると言えます。

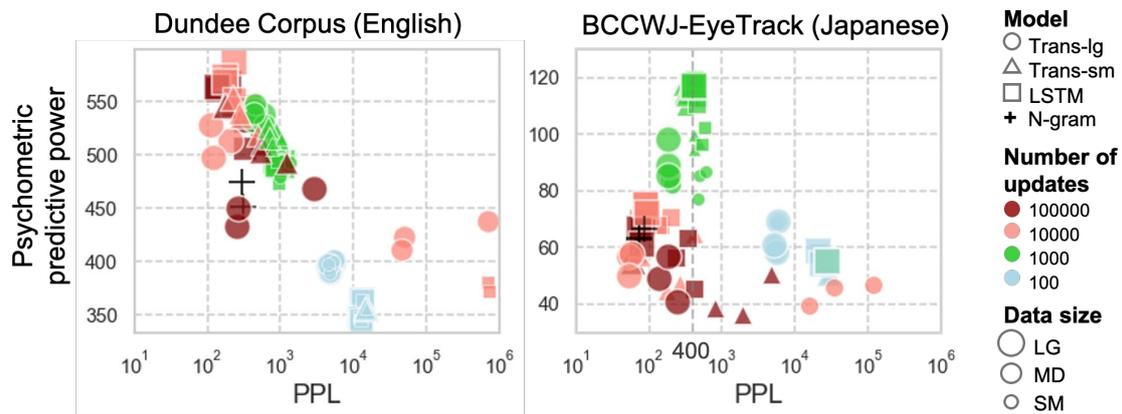


図 2: AI の工学的性能 (x 軸) と認知的妥当性 (y 軸) の関係について、英語における結果 (左図) と日本語における (右図) 結果の比較。各点が異なる条件で訓練した AI に対応。x 軸 (工学的性能) は、値が小さいほど性能が良いことを示している。日本語の結果 (右図) では工学的性能と認知的妥当性に単調な関係がない。

### 【論文情報】

タイトル: Lower Perplexity is Not Always Human-Like 著者: Tatsuki Kuribayashi (東北大学/Langsmith 株式会社), Yohei Oseki (東京大学/RIKEN), Takumi Ito (東北大学/Langsmith 株式会社), Ryo Yoshida (東京大学), Masayuki Asahara (国立国語研究所), Kentaro Inui (東北大学/RIKEN)  
 掲載誌: Proceedings of the Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2021)

URL: <https://aclanthology.org/2021.acl-long.405/>

**【問い合わせ先】**

<研究内容に関する事>

東北大学大学院情報科学研究科

教授 乾健太郎

電話：022-795-7091

E-mail：inui@tohoku.ac.jp

<広報に関する事>

東北大学 大学院情報科学研究科

広報室 佐藤みどり

電話：022-795-4529

E-mail：koho@is.tohoku.ac.jp