

信頼される AI の基盤技術  
2021 年度採択研究者

2021 年度 年次報告書
------------------

大関 洋平

東京大学 大学院総合文化研究科  
講師

認知・脳情報処理による人間らしい言語処理モデルの開発

## § 1. 研究成果の概要

本研究では、自然知能研究としての言語の認知・脳科学と人工知能研究としての自然言語処理を融合することで、認知・脳情報処理に学んだ、人間らしい言語処理モデルを開発することを目的とします。具体的には、①認知処理ベンチマークの構築、②脳情報処理ベンチマークの構築、③人間らしい言語処理モデルの開発、④人間らしい言語処理モデルの評価という4つの研究項目を設定しています。

まず、認知処理ベンチマークの構築について、理論言語学のジャーナルから抽出した人間の容認性判断に関する認知処理ベンチマークである JCoLA を構築しました。研究成果は、言語処理学会第 28 回年次大会で発表し、日本語の言語理解ベンチマークである JGLUE のタスクとして公開する予定です。また、脳情報処理ベンチマークの構築について、東北大学との共同研究として、国立国語研究所の BCCWJ に対して fMRI をアノテーションした人間の脳活動に関する脳情報処理ベンチマークである BCCWJ-fMRI を構築しました。更に、BCCWJ に対して MEG をアノテーションした脳情報処理ベンチマークである BCCWJ-MEG を構築する予定です。

次に、人間らしい言語処理モデルの開発について、記号処理と深層学習を融合した言語処理モデルである再帰的ニューラルネットワーク文法(RNNG)の亜種およびトランスフォーマー文法(TG)を開発しました。研究成果は、言語処理学会第 28 回年次大会で発表し、前者は委員特別賞を受賞しました。加えて、人間らしい言語処理モデルの評価について、トランスフォーマーおよび再帰的ニューラルネットワーク文法を人間の容認性判断および眼球運動に関する既存の認知処理ベンチマークで評価しました。研究成果は、言語処理学会第 28 回年次大会で発表し、計算言語学・自然言語処理のトップカンファレンスである SCiL および EMNLP で発表・出版しました。

### 【代表的な原著論文情報】

- 1) 染谷大河, 大関洋平. 2022. 「日本語版 CoLA の構築」. 言語処理学会第 28 回年次大会, 1872-1877.
- 2) 吉田遼, 大関洋平. 2022. 「トランスフォーマー文法」. 言語処理学会第 28 回年次大会, 448-452.
- 3) Ryo Yoshida, Yohei Oseki. 2021. Learning Argument Structures with Recurrent Neural Network Grammars. Proceedings of the Society for Computation in Linguistics (SCiL), 101-111.
- 4) Ryo Yoshida, Hiroshi Noji, Yohei Oseki. 2021. Modeling Human Sentence Processing with Left-Corner Recurrent Neural Network Grammars. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP), Short Paper, 2964-2973.
- 5) Shingo Shimoda, Lorenzo Jamone, Dimitri Ognibene, Takayuki Nagai, Alessandra Sciutti, Alvaro Costa-Garcia, Yohei Oseki, Tadahiro Taniguchi. 2022. What is the role of the next generation of cognitive robotics? Advanced Robotics 36, 3-16.