

BCCWJ-Brain：言語モデル評価のための脳情報処理ベンチマーク

杉本侑嗣¹ 吉田遼² 浅原正幸³ 鄭 嬌婷⁴

菅野 彰剛⁴ 小泉政利⁴ 大関洋平²

¹大阪大学 ²東京大学

³国立国語研究所 ⁴東北大学

sugimoto.yushi.hmt@osaka-u.ac.jp

{yoshiryo0617,oseki}@g.ecc.u-tokyo.ac.jp

masayu-a@ninja.l.ac.jp

{jeong,akitake.kanno.c8,koizumi}@tohoku.ac.jp

概要

近年自然言語処理と神経科学の融合研究が発展してきており、脳活動データセットの整備が急速にされてきている。すでに一般公開されているデータセットが多数存在するが、得られる脳活動データの性質は、脳機能計測機器に依存する。本研究では、同じ刺激を用いて複数の性質の異なる脳機能計測器から収集された脳情報処理ベンチマークである BCCWJ-Brain を構築し、言語モデルの評価を行った。空間分解度の高いデータ (fMRI) と時間分解度の高いデータ (MEG/EEG) を併用することで、高速に進行する文処理を多面的に検討できる可能性を示した。

1 はじめに

近年の神経科学の領域では、大規模言語モデルなどを用いた認知モデリングが盛んに行われている [1, 2]。モデルの算出する特徴量を用い、モデルと脳活動を対照することで両領域に対する貢献がもたらされることが期待されている。これらの研究において、近年自然な刺激を用いた脳活動データの一般公開が急速に普及しており、研究者各自が脳活動データを説明するモデルを準備し、解析することが可能になってきている。このようなデータは多くの場合、1 種類の脳活動データを用いているが、脳機能計測器ごとに特性が異なる。例えば、機能的磁気共鳴画像法 (functional Magnetic Resonance Imaging, fMRI) は時間分解度が低い一方で空間分解度が高いため、関心のある認知活動の行われている脳活動部位の特定に向いている。他方で脳波 (Electroencephalography,

EEG) は、空間分解度が低い反面、時間分解度は高いため、認知活動の動的な時間変化を捉えるのに適している。脳磁図 (Magnetoencephalography, MEG) は両者の分解度が高いとされている。このように、扱う脳活動データによって捉えられる情報が限られるため、1 種類の脳活動データを用いた場合、検証対象となるモデルの認知的妥当性の包括的な評価ができていないとは言い難い。これらを踏まえ、本研究では、同じ刺激文を用いた複数の脳活動データを取得し、言語モデルの総合的な評価を行うことを試みる。

本研究では、日本語の新聞記事を読んでいる際の脳活動を異なる脳機能計測器でそれぞれ個別にデータ収集した。解析の一例として言語の階層構造を考慮した言語モデルである再帰的ニューラルネットワーク文法 (Recurrent Neural Network Grammar, RNNG) [3] を用いて、fMRI/MEG/EEG の 3 つの脳活動データ解析を行った。先行研究では、英語の fMRI/EEG データで、RNNG の認知的妥当性が検証されている [4, 5]。本研究では、日本語の複数の脳活動データを用いることで、RNNG の認知的妥当性を統一的・包括的に検証することを目標とする。解析の結果、文処理に関連する脳領域、時間幅のそれぞれの観点から RNNG、特に左隅型解析戦略を用いた RNNG の認知的妥当性が示された。

2 脳活動計測・実験設定

脳活動計測には、fMRI/EEG/MEG の実験ですべて同じ刺激文を用い、それぞれ別々の参加者で行った。本研究で得られた脳活動データの詳細な情報は表 1

にまとめた。¹⁾

2.1 実験手続き

刺激文と課題 使用した刺激文は、「現代日本語書き言葉均衡コーパス (Balanced Corpus of Contemporary Written Japanese: BCCWJ, [6])」に収録されている新聞記事 20 件から成り、各記事が文節 (長単位) ごとに視覚提示された。fMRI/MEG/EEG 実験の各実験参加者ごとに提示される記事の順番はランダム化され、各記事の終わりには理解度を確認するために質問が設けられた。本実験では、日本語母語話者 (全員右利き、全員正常な視覚 (矯正を含む)) が、日本語の新聞記事を読んでいる際の脳活動データを取得した (以下、このデータセットを BCCWJ-Brain と呼ぶ)。

2.2 脳活動データのモデリング

本研究では、言語モデルを用いて複数の脳活動データ解析を行った。[7] では、再帰的ニューラルネットワーク文法を用いた BCCWJ-fMRI データ解析が行われたが、本研究では、この言語モデルも含め、同じ刺激文を用いた複数の脳活動データの解析をすることを目的とする。

2.3 言語モデル

LSTM RNN に基づく言語モデル [8]。先行研究では、統語的な依存関係 [9] を扱え、脳活動データへの相関も高いとされる [10]。

RNNG RNN に基づき、自然言語の階層構造を明示的に扱う言語モデル [3]。文処理に関わる脳領域・事象関連電位で LSTM よりも説明力が高いことが先行研究で知られている [4, 5]。[7] では、トップダウン型解析戦略の RNNG と左隅型解析戦略の RNNG を用い、日本語の fMRI データに対して後者の RNNG が前者よりも fMRI データをより良く説明することが示されている。RNNG の推論には、先行研究に基づき、word-synchronous beam search [11] を用いた。²⁾

2.4 評価方法

言語モデルに基づく指標 BCCWJ の新聞記事の文節ごとに各言語モデルが算出する確率に基づいて、認知モデリングを行った。人間の逐次的な文処理において、次の単語を予測しながら高速の処理

をすることは知られており、例えば、ある文脈における次の単語の予測しづらさを指標とするサプライザル ($-\log p(\text{文節}|\text{文脈})$, [13, 14]) は人間の脳活動の負荷を捉えられることが先行研究より知られている [15, 16, 17]。また、RNNG で用いられる指標として distance も用いる。先行研究では、この指標は、RNNG が次の統語的アクションに対する統語的作業量 (syntactic work) を定量化したものとされており、英語の fMRI/EEG データで検証されてきた [4, 5]。

関心領域 本研究では、先行研究ですでに検証されてきた文処理に関連する脳活動部位と事象関連電位に解析範囲を絞った。fMRI/MEG の文処理に関連する脳領域としては、左脳の下前頭回 (Inferior Frontal Gyrus, IFG), 前側頭葉 (Anterior Temporal Lobe, ATL), そして、後部側頭葉 (Posterior Temporal Lobe, PTL) が知られている [18, 19, 20] (各関心領域は、[18] に基づき定義した)。MEG データに関しては、dSPM [21] を用いて信号源推定を行った脳活動データに基づき、図 1 にある各領域ごとの刺激提示から 300-500ms, 400-600ms, 500-700ms の区間のデータを平均化し、解析に用いた。

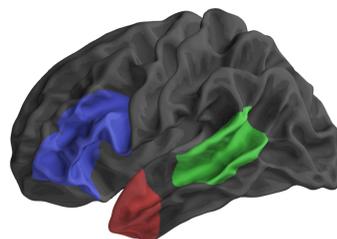


図 1 fMRI/MEG の関心領域: 青 (IFG), 赤 (ATL), 緑 (PTL)。

EEG の関心領域として、LAN, N400 と P600 が文処理に関連する事象関連電位として知られている。LAN は、形態統語的処理や統語的予測に関連しており [22, 23], N400 は、単語の予測困難さに関わる事象関連電位として [24, 25], P600 は、統語・意味情報の統合に関わる事象関連電位として知られている [26, 27, 28]。解析には、図 2 に示されている各センサー内の各時間幅で平均化された脳波データを用いた。

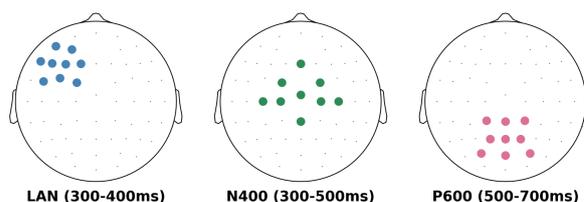


図 2 EEG の関心領域

1) データの前処理に関しては、付録 A を参照されたい。

2) モデル学習等の詳細は [12] を参照されたい。

表1 BCCWJ-Brain の情報

	実験参加者			脳活動データ		刺激文	
	人数	平均年齢		空間分解度	時間分解度	全文節	総文数
fMRI dataset	42 人	21.1	(SD=1.7)	3mm ³	2,000ms	1,642	229
MEG dataset	41 人	21.56	(SD=2.59)	200 チャンネル	1ms (1,000Hz)	1,642	229
EEG dataset	42 人	20.5	(SD=2.89)	64 電極	1ms (1,000Hz)	1,642	229

評価指標 言語モデルから算出されたサプライザルや distance がどれだけ脳活動データを説明するかを検証するため、線形混合回帰モデルを用いて分析した。先行研究に基づき、文節の長さや頻度などの指標を考慮した上で、本研究で調査する言語モデルの指標がどれだけ脳活動データの説明に寄与するか検証する。特に、言語モデルの指標を含めない回帰モデルと含めた回帰モデルの間の対数尤度の差 (Psychometric Predictive Power, PPP) を各言語モデルの Perplexity とともに評価する。回帰モデルとしては、以下の線形混合回帰モデルを用いた³⁾：

$$ROI \sim \text{word_length} + \text{word_frequency} \\ + \text{sentence_ID} + \text{bunsetsu_position} \\ + \{\text{surprisal}, \text{distance}\} \\ + (1|\text{subject}) + (1|\text{article})$$

fMRI データ解析には Nilearn (v0.11.1) [29]⁴⁾ を、MEG/EEG データ解析には、MNE-python (v1.9.0) [30]⁵⁾/Eelbrain (v0.40.3) [31]⁶⁾ をそれぞれ使用し、PPP の算出には、R (v4.4.1) [32] を用いた。

3 結果

図 3 にそれぞれ、MEG, fMRI, EEG 解析の結果がまとめられている。各データセットの各モデルから得られた PPP のベストスコアは付録 B を参照されたい。

3.1 脳活動領域におけるモデル評価

本研究で用いた RNNG の指標の一つである distance が捉える脳活動部位として ATL が挙げられる。MEG の解析の結果より、ATL での活動は初期の時間幅 (300-500ms) で増加する傾向があり、以降徐々に PPP は減少している。傾向として、単語ビーム幅が小さい左隅型解析戦略の RNNG が MEG データの中で一番高い PPP を示しており、先行研究 [12] とも

合致する。fMRI データにおいても、ATL と PTL の領域において同様の傾向を見ることができる。IFG では、トップダウン型解析戦略の RNNG の distance が最も高い PPP を得ており、文処理において、より構造を予測する統語処理は、IFG で行われ、文の逐次的な処理は比較的初期の時間幅の ATL や PTL で行われていることが示唆される。

3.2 時間幅におけるモデル評価

EEG の結果より、LAN においては左隅型解析戦略の RNNG のサプライザルの PPP の増量が顕著である。逐次的な文処理における次の単語・統語アクションの予測によるものであると示唆され、先行研究における LAN の役割と一致する。P600 では、左隅型解析戦略の RNNG の distance がトップダウン型解析戦略の RNNG の distance に比べ PPP が高く、逐次的な文処理に基づいて情報の統合が行われるとされていると言える。N400 の結果では、階層的な構造を考慮しない LSTM から得られたサプライザルの PPP が一番高く、単語の予測困難さという N400 の成分の性質と合致する。

4 おわりに

本研究では、性質の異なる複数の脳活動データを同じ刺激文で収集した BCCWJ-Brain を用いて言語モデルを検証することで、言語モデルの解釈性を高める方向性を示した。従来の研究では、脳活動領域と言語モデルの相関や言語モデルの指標が捉える時間軸のどちらかに注目するのみであった。今回、新たに収集した複数の脳活動データを今後一般公開し、様々な言語モデルの評価に用いることで解釈性の高い言語モデル構築の一端を担うことを期待する。

謝辞

本研究は、JSPS 科研費 JP24H00087, JP25K16284, JST さきがけ JPMJPR21C2, JST CREST JPMJCR2565, JST BOOST JPMJBY24B2 の支援を受けたものです。

3) fMRI データに関しては、参加者ごとに fMRI データ取得時の頭の動きの情報も回帰モデルに追加した。

4) <https://nilearn.github.io/stable/index.html>

5) <https://zenodo.org/records/17675410>

6) <https://eelbrain.readthedocs.io/en/stable/>

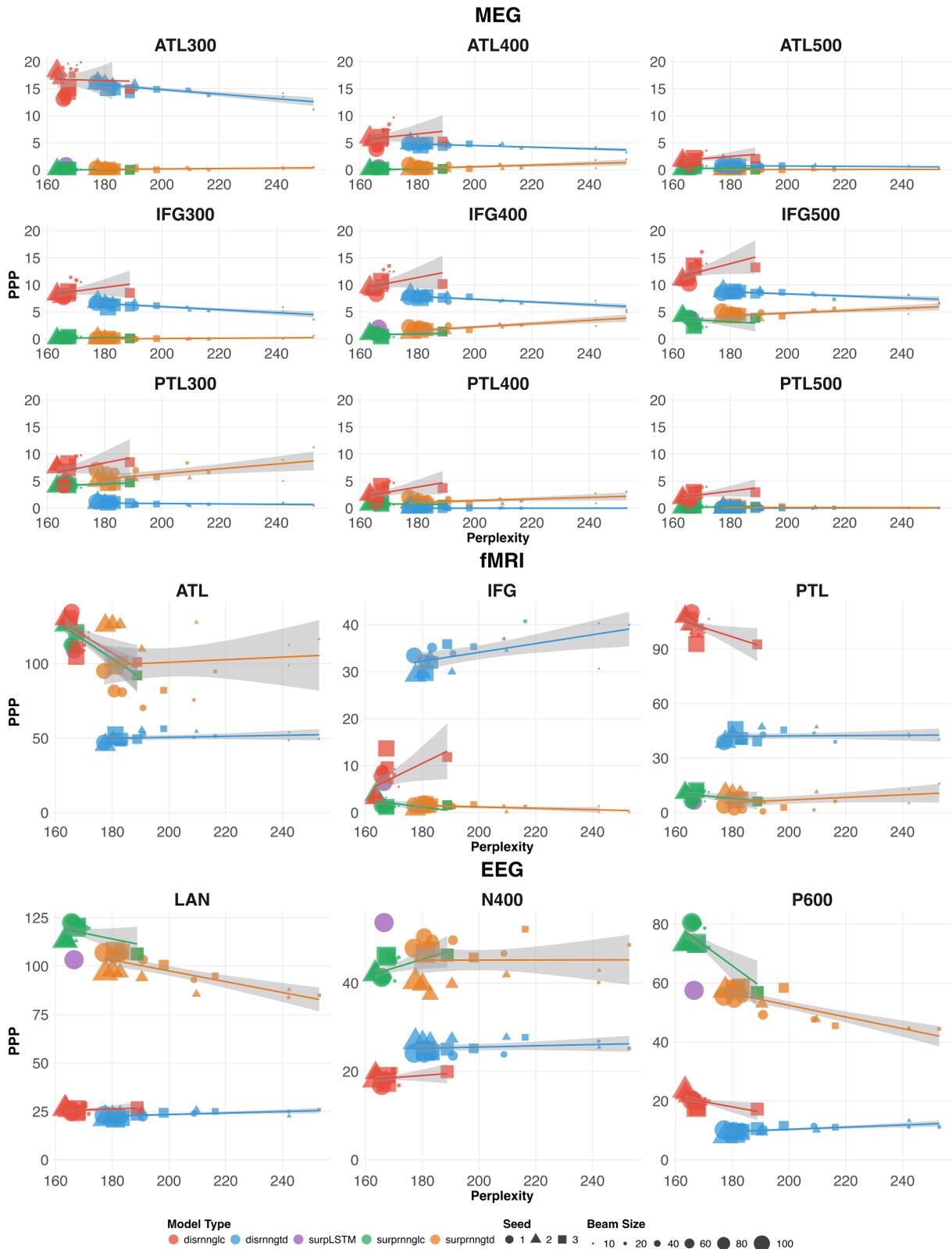


図 3 言語モデルの Perplexity と PPP の関係: Model Type の dis は distance, surp は、サプライザルを表す。rnglc は、左隅型解析戦略の RNNG, rngtd は、トップダウン型解析戦略の RNNG を表す。また、Beam Size は、RNNG の単語ビーム幅を示す。MEG の時間幅の表示は、例えば、ATL300 の場合、ATL の脳領域における 300ms から 500ms の時間幅を示す。よって、ATL のこの時間幅のデータを平均化したものにもとづいて PPP を算出したことになる (PTL, IFG も同様)。EEG の各事象関連電位の時間幅は、図 2 を参照。

参考文献

- [1] Martin Schrimpf, Idan Asher Blank, Greta Tuckute, Carina Kauf, Eghbal A. Hosseini, Nancy Kanwisher, Joshua B. Tenenbaum, and Evelina Fedorenko. The neural architecture of language: Integrative modeling converges on predictive processing. **Proceedings of the National Academy of Sciences**, Vol. 118, No. 45, p. e2105646118, 2021.
- [2] Charlotte Caucheteux and Jean-Rémi King. Brains and algorithms partially converge in natural language processing. **Communications Biology**, Vol. 5, No. 1, p. 134, 2022.
- [3] Chris Dyer, Adhiguna Kuncoro, Miguel Ballesteros, and Noah A. Smith. Recurrent neural network grammars. In **Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, pp. 199–209, San Diego, California, June 2016. Association for Computational Linguistics.
- [4] John Hale, Chris Dyer, Adhiguna Kuncoro, and Jonathan Brennan. Finding syntax in human encephalography with beam search. In Iryna Gurevych and Yusuke Miyao, editors, **Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 2727–2736, Melbourne, Australia, July 2018. Association for Computational Linguistics.
- [5] Jonathan R. Brennan, Chris Dyer, Adhiguna Kuncoro, and John T. Hale. Localizing syntactic predictions using recurrent neural network grammars. **Neuropsychologia**, Vol. 146, p. 107479, 2020.
- [6] Kikuo Maekawa, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka, and Yasuharu Den. Balanced corpus of contemporary written Japanese. **Language Resources and Evaluation**, Vol. 48, No. 2, pp. 345–371, 2014.
- [7] Yushi Sugimoto, Ryo Yoshida, Hyeonjeong Jeong, Masatoshi Koizumi, Jonathan R. Brennan, and Yohei Oseki. Localizing syntactic composition with left-corner recurrent neural network grammars. **Neurobiology of Language**, Vol. 5, No. 1, pp. 201–224, 04 2024.
- [8] Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. **Neural Computation**, Vol. 9, No. 8, pp. 1735–1780, 11 1997.
- [9] Tal Linzen, Emmanuel Dupoux, and Yoav Goldberg. Assessing the ability of LSTMs to learn syntax-sensitive dependencies. **Transactions of the Association for Computational Linguistics**, Vol. 4, pp. 521–535, 2016.
- [10] Shailee Jain and Alexander Huth. Incorporating context into language encoding models for fmri. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, **Advances in Neural Information Processing Systems**, Vol. 31. Curran Associates, Inc., 2018.
- [11] Mitchell Stern, Daniel Fried, and Dan Klein. Effective inference for generative neural parsing. In Martha Palmer, Rebecca Hwa, and Sebastian Riedel, editors, **Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing**, pp. 1695–1700, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.
- [12] Ryo Yoshida, Hiroshi Noji, and Yohei Oseki. Modeling human sentence processing with left-corner recurrent neural network grammars. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing**, pp. 2964–2973, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics.
- [13] John Hale. A probabilistic Earley parser as a psycholinguistic model. In **Second Meeting of the North American Chapter of the Association for Computational Linguistics**, 2001.
- [14] John Hale. Information-theoretical complexity metrics. **Language and Linguistics Compass**, Vol. 10, No. 9, pp. 397–412, 2016.
- [15] Stefan L. Frank, Leun J. Otten, Giulia Galli, and Gabriella Vigliocco. The erp response to the amount of information conveyed by words in sentences. **Brain and Language**, Vol. 140, pp. 1–11, 2015.
- [16] Jonathan R. Brennan and John T. Hale. Hierarchical structure guides rapid linguistic predictions during naturalistic listening. **PLOS ONE**, Vol. 14, No. 1, pp. 1–17, 01 2019.
- [17] Cory Shain, Idan A. Blank, Evelina Fedorenko, Edward Gibson, and William Schuler. Robust effects of working memory demand during naturalistic language comprehension in language-selective cortex. **Journal of Neuroscience**, Vol. 42, No. 39, pp. 7412–7430, 2022.
- [18] Christophe Pallier, Anne-Dominique Devauchelle, and Stanislas Dehaene. Cortical representation of the constituent structure of sentences. **Proceedings of the National Academy of Sciences**, Vol. 108, No. 6, pp. 2522–2527, 2011.
- [19] Douglas K. Bemis and Liina Pykkänen. Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. **Journal of Neuroscience**, Vol. 31, No. 8, pp. 2801–2814, 2011.
- [20] William Matchin, Christian Brodbeck, Christopher Hammerly, and Ellen Lau. The temporal dynamics of structure and content in sentence comprehension: Evidence from fmri-constrained meg. **Human Brain Mapping**, Vol. 40, No. 2, pp. 663–678, 2019.
- [21] Anders M. Dale, Arthur K. Liu, Bruce R. Fischl, Randy L. Buckner, John W. Belliveau, Jeffrey D. Lewine, and Eric Halgren. Dynamic statistical parametric mapping: Combining fmri and meg for high-resolution imaging of cortical activity. **Neuron**, Vol. 26, No. 1, pp. 55–67, 2000.
- [22] Thomas F. Münte, Hans-Jochen Heinze, and George R Mangun. Dissociation of brain activity related to syntactic and semantic aspects of language. **Journal of Cognitive Neuroscience**, Vol. 5, No. 3, pp. 335–344, 07 1993.
- [23] Masataka Yano. Predictive processing of syntactic information: evidence from event-related brain potentials. **Language, Cognition and Neuroscience**, Vol. 33, No. 8, pp. 1017–1031, 2018.
- [24] Marta Kutas and Steven A. Hillyard. Reading senseless sentences: Brain potentials reflect semantic incongruity. **Science**, Vol. 207, No. 4427, pp. 203–205, 1980.
- [25] Katherine A DeLong, Thomas P Urbach, and Marta Kutas. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. **Nature Neuroscience**, Vol. 8, No. 8, pp. 1117–1121, 2005.
- [26] Helen Neville, Janet L. Nicol, Andrew Barss, Kenneth I. Forster, and Merrill F. Garrett. Syntactically based sentence processing classes: Evidence from event-related brain potentials. **Journal of Cognitive Neuroscience**, Vol. 3, No. 2, pp. 151–165, 04 1991.
- [27] Lee Osterhout and Phillip J Holcomb. Event-related brain potentials elicited by syntactic anomaly. **Journal of Memory and Language**, Vol. 31, No. 6, pp. 785–806, 1992.
- [28] Ana C. Gouvea, Colin Phillips, Nina Kazanina, and David Poeppel. The linguistic processes underlying the p600. **Language and Cognitive Processes**, Vol. 25, No. 2, pp. 149–188, 2010.
- [29] Alexandre Abraham, Fabian Pedregosa, Michael Eickenberg, Philippe Gervais, Andreas Mueller, Jean Kossaifi, Alexandre Gramfort, Bertrand Thirion, and Gael Varoquaux. Machine learning for neuroimaging with scikit-learn. **Frontiers in Neuroinformatics**, Vol. Volume 8 - 2014, , 2014.
- [30] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A. Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon Brooks, Lauri Parkkonen, and Matti Hämäläinen. Meg and eeg data analysis with mne-python. **Frontiers in Neuroscience**, Vol. Volume 7 - 2013, , 2013.
- [31] Christian Brodbeck, Proloy Das, Marlies Gillis, Joshua P Kulusingham, Shohini Bhattachali, Phoebe Gaston, Philip Resnik, and Jonathan Z Simon. Eelbrain, a python toolkit for time-continuous analysis with temporal response functions. **eLife**, Vol. 12, p. e85012, nov 2023.
- [32] RStudio Team. **RStudio: Integrated Development Environment for R**. RStudio, PBC., Boston, MA, 2020.

A 前処理

表2 脳活動データの前処理情報

fMRI dataset	MEG dataset	EEG dataset
頭の動き補正	ダウンサンプリング (1,000Hz → 200Hz)	ダウンサンプリング (1,000Hz → 200Hz)
スライス間の撮像時刻のズレの補正	瞬きなどのアーチファクトを独立成分分析 (ICA) を用いて除去	瞬きなどのアーチファクトを独立成分分析 (ICA) を用いて除去
脳解剖画像の脳機能画像への登録	バンドパスフィルタ (0.1-40Hz)	バンドパスフィルタ (0.1-40Hz)
脳解剖画像の分解, 脳機能画像の標準化	-0.1-1,000ms の区間でエポック化し, エポックごとのノイズ除去	-0.1-1,000ms の区間でエポック化し, エポックごとのノイズ除去
脳機能画像の空間平滑化	エポック区間-0.1-0.0ms を使用して, ベースライン補正	全センサーの平均を rereference として, エポック区間-0.1-0.0ms を基準にベースライン補正

B 各脳活動データの各脳領域・各事象関連電位における各言語モデルのベストモデル

表3 MEG データに基づく結果

ATL300				ATL400				ATL500			
model	seed	beam size	PPP	model	seed	beam size	PPP	model	seed	beam size	PPP
surpLSTM	-	-	0.88	surpLSTM	-	-	0.50	surpLSTM	-	-	0.18
surprnngtd	1	10	0.53	surprnngtd	3	10	1.93	surprnngtd	2	60	0.16
surprnnglc	3	80	0.05	surprnnglc	1	100	0.20	surprnnglc	3	20	0.72
disrnngtd	1	100	16.13	disrnngtd	3	80	4.92	disrnngtd	2	20	0.90
disrnnglc	3	10	19.86	disrnnglc	3	10	9.67	disrnnglc	3	10	3.56
IFG300				IFG400				IFG500			
model	seed	beam size	PPP	model	seed	beam size	PPP	model	seed	beam size	PPP
surpLSTM	-	-	0.05	surpLSTM	-	-	2.10	surpLSTM	-	-	3.75
surprnngtd	3	10	0.60	surprnngtd	3	10	4.96	surprnngtd	3	10	6.78
surprnnglc	3	10	0.42	surprnnglc	3	60	1.38	surprnnglc	2	100	4.28
disrnngtd	2	100	6.87	disrnngtd	2	100	8.23	disrnngtd	2	40	9.01
disrnnglc	1	20	11.41	disrnnglc	3	20	13.51	disrnnglc	3	20	16.10
PTL300				PTL400				PTL500			
model	seed	beam size	PPP	model	seed	beam size	PPP	model	seed	beam size	PPP
surpLSTM	-	-	5.02	surpLSTM	-	-	1.25	surpLSTM	-	-	0.03
surprnngtd	3	10	11.24	surprnngtd	3	10	3.07	surprnngtd	3	80	0.32
surprnnglc	3	60	4.75	surprnnglc	1	80	1.09	surprnnglc	3	10	0.36
disrnngtd	1	100	1.19	disrnngtd	2	20	0.03	disrnngtd	3	10	0.10
disrnnglc	3	10	9.79	disrnnglc	3	20	4.94	disrnnglc	3	10	3.93

表4 fMRI データに基づく結果

ATL				IFG				PTL			
model	seed	beam size	PPP	model	seed	beam size	PPP	model	seed	beam size	PPP
surpLSTM	-	-	108.79	surpLSTM	-	-	6.31	surpLSTM	-	-	5.93
surprnngtd	2	20	127.74	surprnngtd	1	80	2.33	surprnngtd	3	10	15.95
surprnnglc	2	10	129.84	surprnnglc	2	100	3.81	surprnnglc	3	100	12.35
disrnngtd	3	40	56.36	disrnngtd	3	20	40.74	disrnngtd	2	20	47.21
disrnnglc	1	100	134.97	disrnnglc	3	100	13.67	disrnnglc	2	10	110.67

表5 EEG データに基づく結果

LAN				N400				P600			
model	seed	beam size	PPP	model	seed	beam size	PPP	model	seed	beam size	PPP
surpLSTM	-	-	103.30	surpLSTM	-	-	53.68	surpLSTM	-	-	57.52
surprnngtd	3	100	107.61	surprnngtd	3	20	52.17	surprnngtd	3	80	58.86
surprnnglc	1	60	122.48	surprnnglc	3	60	46.34	surprnnglc	1	100	80.51
disrnngtd	3	10	25.92	disrnngtd	2	20	27.75	disrnngtd	2	10	13.37
disrnnglc	3	20	28.32	disrnnglc	3	20	20.21	disrnnglc	2	100	23.61