多言語モデルには言語非依存の処理系統が存在するか

手塚陽大 [1]　井之上直也 [1,2]
[1] 北陸先端科学技術大学院大学　[2] 理化学研究所
{hinata-t, naoya-i}@jaist.ac.jp

概要

多言語モデルのプロービングによって言語モデルの多言語性を分析する動きは活発化してきているが，個別言語に依存しない，言語間で共通の処理系統が多言語モデル内に存在するのかについては十分に明らかになっていない．そこで本研究では，第二言語を獲得した言語モデルに焦点を絞り，これらの言語モデルが，第一言語 (L1)，第二言語 (L2) 間で対応する文の意味に関して言語非依存の共通の処理を行っているのかについて，内部表現や発火するニューロンの観点から調査する．実験の結果，言語モデルは，対応する文の意味について個別言語に依存しない共通の処理をしていることが示唆された．

1　はじめに

言語学の第二言語習得論では，個別言語に依存しない言語能力である CUP（**C**ommon **U**nderlying **P**roficiency）の存在を仮定する相互依存仮説が提案されるなど，人間の言語能力の転移について長く議論されてきた [1]．一方，言語処理の分野でも，言語モデルの多言語性や言語能力の転移について調べる動きは活発化しつつある．例えば文献 [2, 3] では，多言語モデル内の言語固有のニューロンを検出し，その多くは最初と最後の数層に位置していることを示している．文献 [4, 5] では，モデルがタスクを解いた時の各言語間のニューロンの重なりを観察している．言語能力の転移については，L1 をいくつかの異なる言語，L2 を英語に設定し，言語モデルの文法性判断ベンチマークにおける L2 の好転移を確認している例がある [6]．さらに文献 [7] では，ある名詞を記憶している知識ニューロンが複数の言語に発火することを示している．

このように，言語モデルにおいて CUP に当たるような能力の存在は示唆されているものの，これが具体的にどのような能力なのかは未だ明らかでない．文献 [7] はこれに対する 1 つの実例を提示しているが，主に名詞を対象とした調査に留まる．さらに，先行実験（付録 A）の結果，英語 (L1) を学習させた言語モデルに別の言語 (L2) を追加学習させることで，英語の文法性判断ベンチマーク BLiMP[8] において L2 学習後に改善する項目があることが分かった．これは言語モデル内部に L1/L2 で共通する処理が存在し，L2 の学習に際して得られた能力を L1 のタスクにおいて活かしている可能性を示唆する．

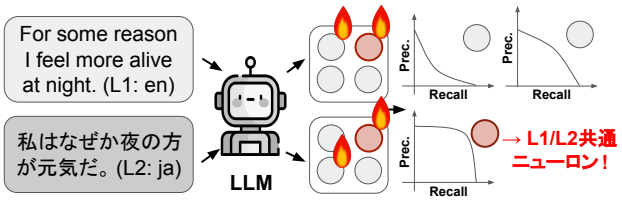


図 1: **L1 / L2 双方に発火する共通ニューロンの検出例.** 実験では，対訳ペア（ラベル 1）と非対訳ペア（ラベル 0）の文対を入力し，L1/L2 共通ニューロンの発火値をもとに PR 曲線の AUC を計算することで，対応する文の意味のみに強い影響を持つ L1/L2 共通ニューロンを検出する．

そこで，本研究では個別言語に依存しない処理が行われ得る情報として文の意味に着目し，**L2 を獲得した言語モデルでは，L1/L2 間で対応する文の意味に関して，共通の処理が行われるか**について調査する．L1/L2 に対象言語を絞るのは，観察のしやすさのためである．実験では，まず内部表現を分析した後 (§2)，より詳細な分析のため，対応する文の意味に特化して発火する L1/L2 共通ニューロンを検出し，発火値を改ざんする介入実験をいくつか行う (§3)．本研究の貢献は次の通りである．

- L1/L2 共通に強く発火し，対応する文の意味にのみ強い影響をもつニューロンの存在を示す．
- 第二言語を獲得した言語モデルが，対応する文の意味（対訳文ペア）について個別言語に依存しない共通の処理をしていることを示唆する．

2　隠れ状態・内部表現の分析

2.1　対応する文の意味には言語に依らず似た内部表現が作られるか

多言語モデル内では，言語に依らず，対応する文の意味（対訳ペア）には類似した内部表現が作られるのかを調査する．この際，最後のトークンのみを観察対象とする．なお，実験では各 L1/L2 ペアに対応した対訳コーパスを使用する．L1/L2 の対訳文ペアと非対訳文ペアをモデルに入力することで，以下の 2 つを実験により確かめる．詳細は付録 4 に記載する．

1. 対訳ペア・非対訳ペア間の隠れ状態の類似度に明確な差が出るか．
2. 入力された L1/L2 隠れ状態が対応する文の意味を表すかどうかを精度良く予測できるか．

2.2　実験結果

対訳ペア・非対訳ペア間で隠れ状態の類似度に明確な差が出るか　図 2 は，L2 を追加学習した LLaMA3 に対訳・非対訳ペア（2000 ずつ）を与えた時の L1/L2 間の隠れ状態のコサイン類似度を，層ごとに平均したものである．明らかに，対訳ペア（青）を与えた時の方が言語に依らず類似度が高く，似た表現を作れていることがわかる[1)]．また，Self-Attention・MLP（**M**ulti **L**ayer **P**erceptron）直後の表現を同じ様に比較したものを付録 B.2 に示す．

分類モデルは入力された L1/L2 隠れ状態が対応する文の意味を表すかどうかを精度良く予測できるか　付録 B.1 に，ロジスティック回帰モデルを構築・テストし，隠れ状態の分類精度を測った結果を示す．全ての層において非常に高い精度で対訳ペアか・非対訳ペアかを分類できている結果となった．

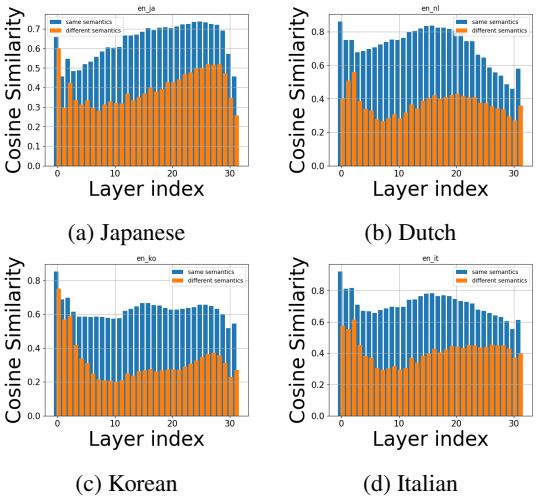
これらの結果から，モデルは L1/L2 間で対応する文の意味に対して，言語に依らず似た内部表現を作れていることがわかった．

3　ニューロンの検出と制御

3.1　対応する文の意味に特化した L1/L2 共通ニューロンは存在するか

クロスリンガルに発火する知識ニューロンが多言語モデルの MLP に存在することを確認している

1)　異なる意味表現のコサイン類似度もある程度高いことに関しては，異方性 (Anisotropy) の影響が考えられる [9, 10].



(a) Japanese　(b) Dutch

(c) Korean　(d) Italian

図 2: **L2 を追加学習した LLaMA3 モデルにおける，L1/L2 文ペアを与えた時の各層の隠れ状態の類似度**
L1/L2 の対訳（青）・非対訳（オレンジ）2000 ペアずつを入力した時の，隠れ状態の類似度を層ごとに平均した．

文献 [7] の結果を受けて，我々は，**MLP 内の L1/L2 双方に発火する共通ニューロンにて，言語非依存の処理が行われる**と仮定する．この仮説に基づいて，L1/L2 双方によく発火し，かつ対応する文の意味にのみ強く反応するニューロンが MLP に存在するかどうかを調べる．実験では，まず §3.3 に説明する方法でこのようなニューロンの検出を試みる．その後，検出したニューロンの発火値改ざんし，以下の 2 つを確認する介入実験を行う．

1. §2 の実験（隠れ状態の類似度の測定）を再度行い，検出したニューロンが本当に対応する文の意味のみに強い影響を持つことを確認する．
2. 1 を確認した上で，先行実験（付録 A）で示した文法性判断タスクを再度モデルに解かせ，当該タスクにおける検出したニューロンの影響を確認する．

なお，実験設定は §2 を踏襲する．

3.2　準備：ニューロン・活性化値の定義

実験に先立ち，ニューロン及び活性化値の定義を明確化する．ニューロンを検出している先行研究はいくつかあり [2, 3, 4, 5, 11]，MLP については活性化関数の出力をニューロンの単位および活性化値としている研究が多い．しかし我々は，以下の理由から，LLaMA3 と GPT2 でそれぞれ別の場所をニューロンおよび活性化値として定義する．

LLaMA3 のニューロン　LLaMA3 では，MLP は以下の式で計算される．

$$\mathrm{MLP}^{(l)}(X') = (a(X'M_I^{(l)}) \odot X'M_G^{(l)})M_O^{(l)} \quad (1)$$

l, a, $M_I^{(l)}$, $M_G^{(l)}$, $M_O^{(l)}$ はぞれぞれ順に，layer index，活性化関数，gate projection，up projection，down projection を指す．式 1 からわかるように，LLaMA3 の MLP では活性化関数の出力と up projection の出力のアダマール積をとるため，仮に活性化関数 a の出力が大きい値でも，up projection の出力の影響次第では活性化値が小さくなり，結果的に次のニューロンへの影響が弱まる場合がある．これを踏まえて我々は，down projection への入力となる以下の部分をニューロンおよび活性化値として定義する．

$$\mathrm{Neurons}^{(l)} = a(X'M_I^{(l)}) \odot X'M_G^{(l)} \quad (2)$$

GPT2 のニューロン　GPT2 の MLP は以下の式 3 で計算されるため，単に活性化関数の出力をニューロンおよび活性化値とする．

$$\mathrm{MLP}^{(l)}(X') = a(X'M_I^{(l)})M_O^{(l)} \quad (3)$$

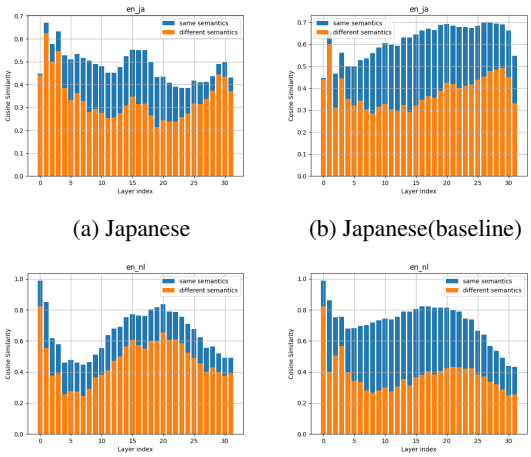
3.3　対応する文の意味にのみ強い影響を持つ L1/L2 共通ニューロンの検出方法

L1/L2 双方に強く発火し，かつ対応する文の意味（対訳ペア）にのみ強い影響を持つニューロンの検出を試みる．ここでの我々の目的は，**対訳ペアには強い発火をし，非対訳ペアには非常に弱い，もしくは発火しない L1/L2 共通ニューロンを検出すること**である．そのために，[11] の手法を部分的に踏襲する（図 1 に検出方法の概要を示す）．具体的には，各 L1/L2 共通ニューロンの活性化値をそのニューロンの予測値とみなし，L1/L2 文ペアの入力が，対応する文の意味（対訳ペア: ラベル 1）を表すか・異なる意味（非対訳ペア: ラベル 0) を表すかを予測する以下のような分類問題として形式化する．

$$f(S_{L1}, S_{L2}) = \begin{cases} 1 & \text{if } \left(|\alpha_i^l(S_{L1})| + |\alpha_i^l(S_{L2})|\right)/2 \geq \tau \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

S_{L1}/S_{L2}, α_i^l, τ はそれぞれ，L1/L2 の文ペア，l 層目・i 番目のニューロンの活性化値を返す関数，閾値を指す．[2)] 予測結果をもとに，PR 曲線の AUC（**A**rea **U**nder the **C**urve）である平均適合率（以下，AP）を計算する．そして算出した AP をもとに各 L1/L2 共通ニューロンを降順に並べ，上位 n 件を対象のニューロンとする．

2)　式 4 で 2 つの活性化値 (L1/L2) の絶対値の平均を取っている意図は，(1) 負の強い発火も考慮するため (2) L1/L2 どちらにも強い発火をしているニューロンが上位にくるようにするためである．



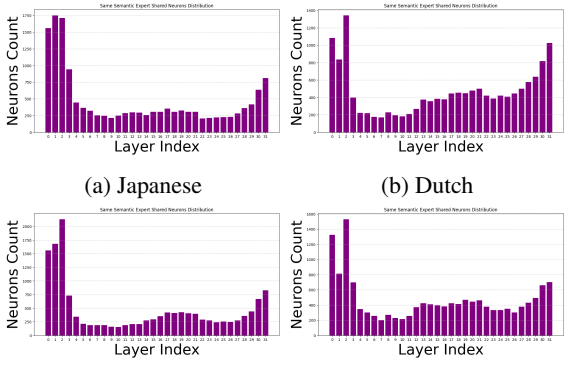
(a) Japanese　(b) Japanese(baseline)

(c) Dutch　(d) Dutch(baseline)

図 3: **AP 上位 15000 個（全体の約 3.3%）の L1/L2 共通ニューロンを無効化したときの隠れ状態の類似度.** 右（baseline）は，比較対象として L1/L2 共通ニューロンのうち，AP 上位 15000 個の補集合から 15000 個をランダムに抽出し，同様に無効化したもの.

3.4　実験結果

検出したニューロンは確かに対応する意味表現のみに強い影響をもつ　図 3 は，§3.3 の方法によって検出したニューロンのうち，AP 値が上位 15000 件（全体の約 3.3%）の発火値を 0 に改ざんし，当該ニューロンの影響をほぼ無効化した上で再度隠れ状態の類似度を測ったものである．青（対訳ペア）の類似度は大きく下がっているのに対し，オレンジ（非対訳ペア）の類似度はほとんど下がっておらず，両者の類似度の差が著しく縮まっている．付録 B.3 にも示す通り，この傾向は英語に近い言語の方が顕著であった．したがって，対応する文の意味のみに強い発火をする L1/L2 共通ニューロンが存在し，それらを適切に検出できていることがわかる．図 4 に検出したニューロンの分布を示すが，AP 値が高いニューロンほど最初と最後の数層に多く位置している傾向が観察された．興味深いことに，これは言語固有ニューロンが多くある層と一致する [2, 3]．さらに，付録 B.2 に示すとおり，Self-Attention 直後に比べて MLP 直後の内部表現の方が，最初と最後の数層において対訳・非対訳をよく区別できていることから，検出したニューロンが貢献していることが



(a) Japanese (b) Dutch

(c) Korean (d) Italian

図 4: **LLaMA3 における共通の意味表現に特化した L1/L2 共通ニューロンの分布（上位 15000 件: 全体の約 3.3%）**
横軸は layer index，縦軸は総ニューロン数.

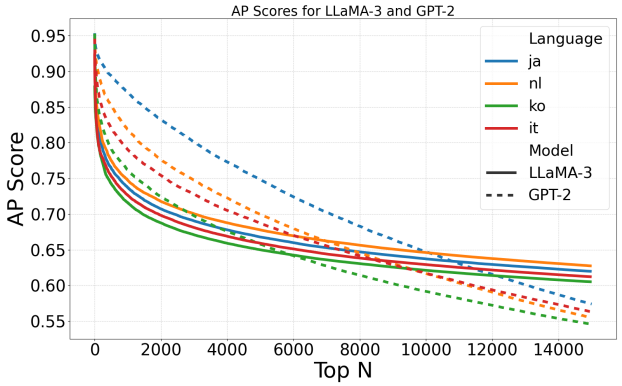


図 5: **検出したニューロンのうち上位 15000 件の AP 値.**
実線は LLaMA3，点線は GPT2．横軸は上位 n 件（左にいくほど上位），縦軸は AP 値．モデルや L2 を問わず，上位は AP 値が非常に高いことがわかる.

示唆される．なお，改ざんする数をさらに増やしていくと，青（対訳）・オレンジ（非対訳）の差はさらに縮まるとともに，オレンジ（非対訳ペアの類似度）が上昇する傾向が観察された．以上より，

- L1/L2 共通に強く発火し，かつ対応する文の意味のみに強い影響をもつニューロンが MLP に存在する
- 上位の AP 値（図 5）が非常に高いことから，上位のニューロンの活性化値を見れば，異なる言語の文対が同じ意味かどうか精度良く分かる

ということがわかった[3)].

検出したニューロンは文法性判断タスクにおいてどの程度影響をもつか §3.3 の方法によって検出したニューロンが，文法性判断タスクにおいて強い影響をもつのか，また，先行実験（付録 A）で示した精度が向上した結果と因果関係があるのかについて調査する．L2（日本語）を追加学習した LLaMA3 において，検出したニューロンの上位 10000 件（全体の約 2%）の発火値を 0 に改ざんした上で，再度文法性判断ベンチマーク BLiMP の精度を測定し，元のスコアとの落差を見ることで，前節で検出したニューロンの影響を調べる．また，L2 の文法タスクにおける影響も調べるため，日本語の文法判断性ベンチマークである JBLiMP[12] の精度も同様に測定する．また，前節で検出したニューロン（同じ意味）に加え，比較対象として，異なる意味表現（非対訳ペア）2000 ペアを入力した時に，強く発火した L1/L2 共通ニューロン（違う意味）と，タスクに使用された言語には強く発火したが，もう一方の言語にはほとんど発火しなかったニューロン（タスク言語）のうち，発火値の累計が上位 10000 件ニューロンを同様に無効化した．表 1 から分かる通り，§3.3 で検出したニューロン（同じ意味）に介入しても精度はそこまで落ちず，影響は大きくないことがわかった．しかし，異なる意味表現の文対に強い発火を見せる L1/L2 共通ニューロン（違う意味）の影響は，タスク言語のみに強く発火するニューロンと比べても明らかに大きく，本タスクにおいて L1/L2 共通によく発火するニューロンの影響が大きいことがわかり，L1/L2 に依存しない何らかの処理が重要な働きをしている可能性が高いことが示唆された．この因果関係については今後の調査課題である.

表 1: **L2(ja) 獲得後の LLaMA3 の文法性判断ベンチマークにおける検出したニューロンの影響 (OVERALL)**

タスク	元スコア	同じ意味	違う意味	タスク言語
BLiMP	0.776	0.745	0.508	0.755 (en)
JBLiMP	0.701	0.670	0.575	0.698 (ja)

4 おわりに

本研究では，第二言語 (L2) を獲得した言語モデルにおいて，L1/L2 間で対応する文の意味に関して共通の処理が行われるのかについて調査した．結果として，対訳文ペアと非対訳文ペアを入力した時の内部表現における類似度・分類精度の明確な違いや，L1/L2 共通に強く発火し，かつ対応する文の意味に特化したニューロンの存在が MLP において確認されるなど，言語モデルにおいて，対応する文の意味については個別言語に依存しない形で共通の処理がなされていることが示唆された.

3) L1/L2 間の活性化値の符号の一致を考慮するため，L1 活性化値と L2 活性化値の積 $\alpha(S_{L1}) \cdot \alpha(S_{L2})$ を L1/L2 共通ニューロンの発火値とし，同じように介入実験を行ったが，絶対値の平均をとる式 4 と比べて大きな差は見られなかった.

謝辞

本研究は中島記念国際交流財団の助成を受けたものです．また，JAIST の坂井吉弘氏には，本稿の執筆に際して非常に有益な助言をいただきました．この場を借りて感謝申し上げます．

参考文献

[1] James Cummins. Linguistic interdependence and the educational development of bilingual children. **Review of Educational Research Vol.49, No.2 (Spring, 1979)**, pp. 222–251, 1979.

[2] Takeshi Kojima, Itsuki Okimura, Yusuke Iwasawa, Hitomi Yanaka, and Yutaka Matsuo. On the multilingual ability of decoder-based pre-trained language models: Finding and controlling language-specific neurons. In **Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)**, 2024.

[3] Tianyi Tang, Wenyang Luo, Haoyang Huang, Dongdong Zhang, Xiaolei Wang, Xin Zhao, Furu Wei, and Ji-Rong Wen. Language-specific neurons: The key to multilingual capabilities in large language models. In **Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)**, 2024.

[4] Karolina Stanczak, Edoardo Ponti, Lucas Torroba Hennigen, Ryan Cotterell, and Isabelle Augenstein. Same neurons, different languages: Probing morphosyntax in multilingual pre-trained models. In **Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, 2022.

[5] Weixuan Wang, Barry Haddow, Minghao Wu, Wei Peng, and Alexandra Birch. Sharing matters: Analysing neurons across languages and tasks in llms. **arXiv preprint arXiv:2406.09265**, 2024.

[6] Miyu Oba, Tatsuki Kuribayashi, Hiroki Ouchi, and Taro Watanabe. Second language acquisition of neural language models. In **Findings of the Association for Computational Linguistics: ACL 2023**, 2023.

[7] Yubo Chen Kang Liu Jun Zhao Yuheng Chen, Pengfei Cao. Journey to the center of the knowledge neurons: Discoveries of language-independent knowledge neurons and degenerate knowledge neurons. In **AAAI Conference on Artificial Intelligence**, 2024.

[8] Alex Warstadt, Alicia Parrish, Haokun Liu, Anhad Mohananey, Wei Peng, Sheng-Fu Wang, and Samuel R. Bowman. BLiMP: The benchmark of linguistic minimal pairs for english. **Transactions of the Association for Computational Linguistics**, Vol. 8, pp. 377–392, 2020.

[9] Nathan Godey, Éric Clergerie, and Benoît Sagot. Anisotropy is inherent to self-attention in transformers. In **Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)**, pp. 35–48, St. Julian's, Malta, March 2024. Association for Computational Linguistics.

[10] Kawin Ethayarajh. How contextual are contextualized word representations? Comparing the geometry of BERT, ELMo, and GPT-2 embeddings. In **Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)**, pp. 55–65, Hong Kong, China, November 2019. Association for Computational Linguistics.

[11] Xavier Suau, Luca Zappella, and Nicholas Apostoloff. Self-conditioning pre-trained language models. **International Conference on Machine Learning**, 2022.

[12] Taiga Someya and Yohei Oseki. JBLiMP: Japanese benchmark of linguistic minimal pairs. In **Findings of the Association for Computational Linguistics: EACL 2023**, pp. 1581–1594, Dubrovnik, Croatia, May 2023. Association for Computational Linguistics.

[13] Meta. Introducing meta llama 3: The most capable openly available llm to date, April 2024. https://ai.meta.com/blog/meta-llama-3/.

[14] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. **OpenAI**, 2019.

[15] Naoaki Okazaki, Kakeru Hattori, Hirai Shota, Hiroki Iida, Masanari Ohi, Kazuki Fujii, Taishi Nakamura, Mengsay Loem, Rio Yokota, and Sakae Mizuki. Building a large japanese web corpus for large language models. In **Proceedings of the First Conference on Language Modeling**, COLM, University of Pennsylvania, USA, October 2024.

[16] Kei Sawada, Tianyu Zhao, Makoto Shing, Kentaro Mitsui, Akio Kaga, Yukiya Hono, Toshiaki Wakatsuki, and Koh Mitsuda. Release of pre-trained models for the Japanese language. In **Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)**, pp. 13898–13905, 5 2024.

[17] ReBatch. Llama-3-8b-dutch model card, 2024. https://huggingface.co/ReBatch/Llama-3-8B-dutch.

[18] Wietse de Vries and Malvina Nissim. As good as new. how to successfully recycle english gpt-2 to make models for other languages. **arXiv preprint arXiv:2012.05628**, 2020.

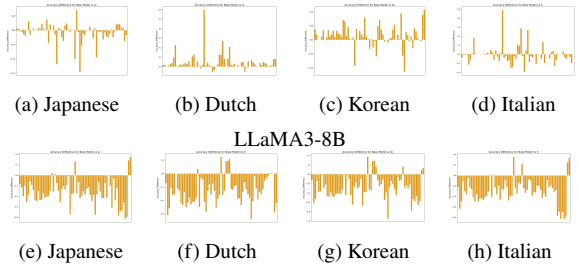
[19] Junbum L. Llama-3-koen model card, 2024. https://huggingface.co/beomi/Llama-3-KoEn-8B.

[20] SKT-AI. Kogpt2, 2020. https://github.com/SKT-AI/KoGPT2.

[21] DeepMount. Llama-3-8b-ita model card, 2024. https://huggingface.co/DeepMount00/Llama-3-8b-Ita.

[22] Jörg Tiedemann. The tatoeba translation challenge – realistic data sets for low resource and multilingual MT. In **Proceedings of the Fifth Conference on Machine Translation**, pp. 1174–1182. Association for Computational Linguistics, November 2020.

使用したモデル・コーパス　LLaMA3-8B，GPT2-small の元のモデルを L1(英語) 学習済みのベースモデル[4]とし，L2（日本語 [15, 16]，オランダ語 [17, 18]，韓国語 [19, 20]，イタリア語 [21, 18]）を追加学習したモデルを使用した．対訳コーパスは tatoeba[22] を採用し，実験全体で対訳文 2000 ペア, 非対訳文 2000 ペアを使用した．



(a) Japanese　(b) Dutch　(c) Korean　(d) Italian

(e) Japanese　(f) Dutch　(g) Korean　(h) Italian

図 6: **L1 獲得後と比較した L2 獲得後の BLiMP の精度**

A　先行実験の示唆 (BLiMP)

図 6 は，L1（英語）獲得後のベースモデルと比較した，L2 獲得後モデルの BLiMP（全 67 項目）の精度である．上に伸びているバーはベースモデルと比較して精度が向上している項目，下に伸びているバーは L2 獲得後に落ちている項目を指す．パラメータ数の多い LLaMA3 の方が向上している項目が多い結果となったが，GPT2 においても L2 獲得後にいくつかの項目で精度が向上していることがわかる．この結果は，L2 の追学習時に得た知識を L1 のタスクにおいて生かせている可能性を示唆する．なお，BLiMP では文法的に容認可能な文と不可能な文のペアが項目ごとに 1000 ペアずつ含まれており，本稿では，文中の各トークンに割り当てられた対数確率の平均をその文の自然な生成確率とし，容認可能な文と不可能な文の各ペアにおいて，前者に高い生成確率が付与された割合をスコアとした．

B　隠れ状態の類似度

B.1　ロジスティック回帰モデルによる隠れ状態の分類

図 7 は，ロジスティック回帰モデルを訓練して構築し，隠れ状態の分類を行った結果である．対訳 (ラベル 1)，非対訳 (ラベル 0) を 2000 ペアずつ用意し，層化 k 分割交差検証 (k=10) を行った際の，テスト事例に対する正解率の平均を層ごとに計算した[5]．モデルを問わず，全ての層で 9 割以上の精度で分類できている結果となった．

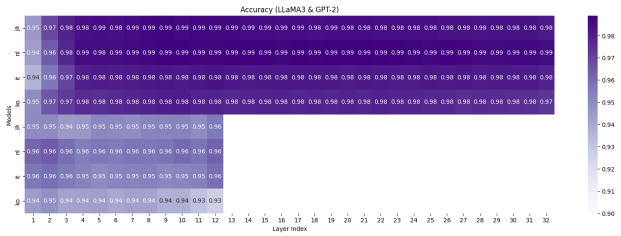
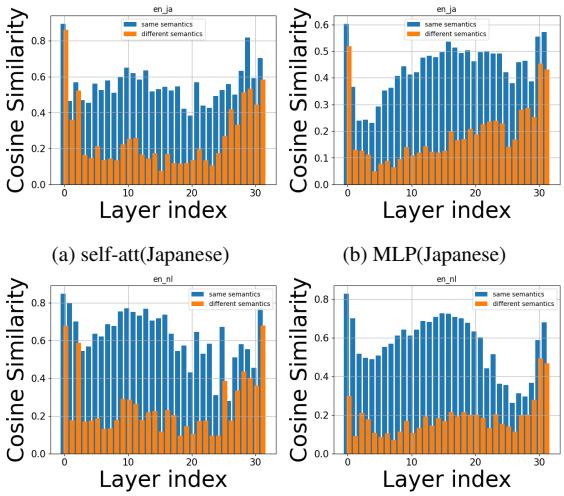


図 7: **ロジスティック回帰モデルによる隠れ状態の分類**

B.2　Self-Attention，MLP 直後の類似度

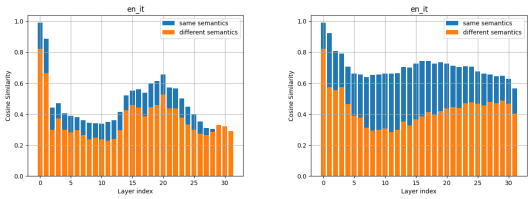


(a) self-att(Japanese)　(b) MLP(Japanese)

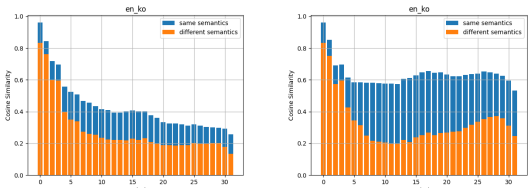
(c) self-att(Dutch)　(d) MLP(Dutch)

図 8: **L2（日本語，オランダ語）を追学習した LLaMA3 における内部表現のコサイン類似度 (青：対訳，オレンジ：非対訳)**

B.3　発火値を改ざんした状態の隠れ状態の類似度



(a) Italian　(b) Italian(baseline)



(c) Korean　(d) Korean(baseline)

図 9: **AP 上位 15000 個（全体の約 3.3%）の L1/L2 共通ニューロンを無効化したときの各層の隠れ状態の類似度.**

4)　両モデルとも，事前学習時のデータはほとんどが英語である [13, 14].

5)　L1 の隠れ状態と L2 の隠れ状態を連結して 1 つの特徴量とした．