

End-to-End モデルに基づく 漸進的係り受け解析と未入力文節主辞予測の同時実行

海野 博揮¹ 大野 誠寛² 伊藤 滉一郎¹ 松原 茂樹^{1,3}

¹ 名古屋大学大学院情報学研究科 ² 東京電機大学大学院未来科学研究科

³ 名古屋大学情報基盤センター

unno.hiroki.t9@s.mail.nagoya-u.ac.jp ohno@mail.dendai.ac.jp

{ito.koichiro.z5,matsubara.shigeki.z8}@f.mail.nagoya-u.ac.jp

概要

同時通訳のようなリアルタイム言語処理システムでは、文を構成する要素が順次入力される状況において、その入力に即応した出力が求められる。このようなシステムで構文情報を利用できるようにするために、漸進的係り受け解析の研究が行われている。本論文では、深層学習モデルを用いて、End-to-End で漸進的係り受け解析を行う手法を提案する。本手法では、未入力の係り先文節の主辞予測を同時に実行し、漸進的係り受け解析の精度向上を図る。実験により提案手法の解析性能を評価した。

1 はじめに

同時通訳 [1] や音声対話システム [2], リアルタイム字幕生成 [3, 4] などのリアルタイム言語処理システムでは、文を構成する要素が順次入力される状況において、その入力に即応した出力が求められる。このようなシステムに対して構文情報を提供するために、漸進的係り受け解析、すなわち、入力途中の段階で文節間の係り受け関係を同定する技術の研究が数多く行われている [5, 6, 7, 8, 9, 10]。その中でも、入力途中の段階でどのような係り受け情報を出力するかについて様々な検討がなされている。

なかでも橋本ら [5] は、係り受け構造を漸進的に解析するとともに、未入力の係り先文節の主辞を予測することを提案しており、入力途中段階で多彩な情報を提供することができる (図 1)。しかしながら、拡張した Shift-Reduce 法において深層学習を限定的に用いているに過ぎず、文内の文脈情報を効果的に利用できていない可能性がある。

本論文では、深層学習モデルを用いて、End-to-End で漸進的係り受け解析を行う手法を提案する。提案

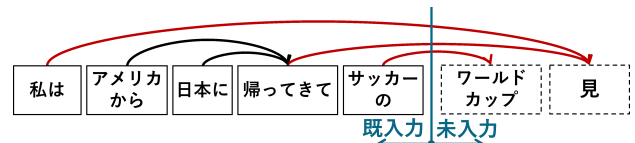


図 1 本研究で漸進的に出力する係り受け構造

手法では、文節が追加されるたびに、図 1 に示す係り受け構造を解析する。また、先行研究 [5] と同様に、未入力の係り先文節の主辞を同時に予測することで解析精度の向上を図る。本研究は End-to-End モデルによる漸進的係り受け解析の可能性を探究するものである。

2 関連研究

漸進的係り受け解析は、入力文の途中で文節間の係り受け関係を同定する技術であり、これまでも多く研究されている [5, 6, 7, 8, 9, 10]。そこでは、入力途中の段階でどのような係り受け情報を出力するかについて様々な検討がなされている。

それらの中でも、橋本ら [5] は、Shift-Reduce アルゴリズム [11] を拡張し、既入力文節列から、図 1 に示す係り受け構造を解析する手法を提案している。この手法は、入力途中段階で多彩な情報を後段のシステムに提供することができる。また、人間が漸進的係り受け解析と未入力文節の予測を同時に行っているという仮定のもと、未入力の係り先文節の主辞を同時に予測することで、係り受け解析の精度を向上させることができることを示している。

一方、近年では、係り受け解析をタグ付けベースの手法として定式化することにより End-to-End で解く手法が提案されている [12, 13, 14, 15, 16, 17, 18, 19, 20]。これらの手法は、係り受け解析を単語レベルのタグ付け問題として定式化している。

また、柴田ら [21] は、日本語における係り受け解

析を head selection 問題 [12] として捉え、形態素と基本句の2つのレベルで係り先を予測する手法を提案している。さらに、吉田ら [22] は、柴田らの手法 [21] を拡張し、係り先が未入力であることを示す特殊トークンを導入し、係り先が未入力と明示した漸進的係り受け解析を行う手法を提案している。

3 提案手法

本研究では、橋本ら [5] と同一の問題設定を採用する。すなわち、文節 n 個からなる文 $b_1 \dots b_n$ に対して、文節 b_t ($1 \leq t < n$) が入力されるたびに、既入力文節列 $B_t = b_1 \dots b_t$ の各文節 b_i ($1 \leq i \leq t$) の係り先（未入力文節との構文的関係も含む）と、その未入力文節の主辞トークンを明示した構造（図 1）を End-to-End で同定する。本手法の概要を図 2 に示す。エンコーダへの入力はいずれも既入力文節列 $B_t = b_1 \dots b_t$ に未入力の係り先文節を示す特殊トークン ([UND]) を追加したものとする。エンコーダの最終層に2つの独立した全結合層を追加する。1つは係り先の文節を解析するための全結合層であり、もう1つは未入力の係り先文節の主辞トークンを予測するための全結合層である。

未入力文節との構文的関係を含む既入力文節列 $B_t = b_1 \dots b_t$ に対するモデルの学習は係り受け解析と主辞予測のマルチタスク学習で行った。損失関数はそれぞれ交差エントロピー損失を用い、全体の損失は係数 λ を用いて以下の式で計算した。

$$Loss = \lambda \times \text{係り受け解析損失} + (1 - \lambda) \times \text{主辞予測損失} \quad (1)$$

3.1 漸進的係り受け解析

提案手法では、先行研究 [21] にない、係り受け解析を head selection 問題 [12] として捉え、各入力文節の主辞トークンに対して係り先となる文節の番号を推定する。この際、入力の末尾に [ROOT] を追加し、1文の根となる文節の主辞トークンの係り先を [ROOT] とする。トークンレベルの解析を通じて文節間の係り受け関係をモデル化する。

係り先が未入力の文節である場合、先行研究 [22] にない、係り先が未入力であることを示す特殊トークン ([UND]) を導入する。この特殊トークンは既入力文節列と [ROOT] の間に追加され、モデルが係り先を未入力として判断した場合に選択される。さらに、提案手法では未入力の係り先を区別するため、複数の [UND] を入力列に追加する。この [UND]

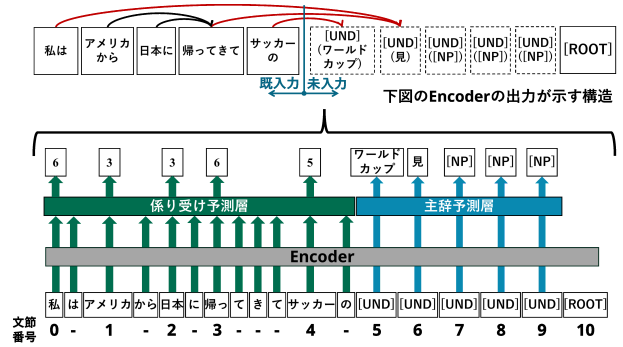


図 2 提案手法の概要

の数は、入力文節が全て異なる未入力文節に係るケースを想定し、入力文節数と同数に設定する。

3.2 未入力の係り先文節の主辞予測

未入力の係り先文節の予測では、各 [UND] に対して各単語確率を計算し、最も確率が高い単語を未入力の係り先文節の主辞¹⁾として選択する。一方で、係り元が存在しない [UND] に対しては、主辞予測を回避するための特殊トークン ([NP]) を導入し、不要な主辞予測を回避する。

4 実験

提案手法の有効性を評価するため、実験を行った。

4.1 実験の概要

実験には、同時通訳データベース (SIDB) [24] に収録されている日本語講演の書き起こしデータを使用した。このデータには、IPA 辞書に基づいた形態素情報、文節境界情報、節境界情報、係り受け情報が人手で付与されている。

実験は、全 16 講演のうち、1 講演をテストデータとし、残りの 15 講演を学習データとする実験を 16 回繰り返す交差検定により実施した。ただし、16 講演のうち、2 講演は開発データとし評価には用いず、残りの 14 講演で評価した。これらは先行研究 [5, 25] と同一の実験設定である。

本実験では、1 文の各文節が入力されるたびに、その既入力文節列に対するデータの作成を繰り返すことで、実験に用いるデータを作成した。そのため、1 文に対して文節数と同数のデータが作成され、作成したデータを全て用いると過学習を起こす可能

1) 主辞の定義は内元らの定義 [23] を参考に、そこから助動詞を除外したものとした。また、非自立語は原則として主辞に含めないが、文節内に他の主辞候補が存在しない場合に限り、主辞として扱った。

表 1 係り受け解析結果 (R:再現率, P:適合率, F:F値)

手法	係り先が既入力			係り先が未入力		
	R	P	F	R	P	F
SR 法	82.70	81.72	82.21	73.58	74.26	73.92
SR 法 ($\lambda = 1$)	82.59	81.62	82.10	73.39	74.03	73.70
提案手法	78.40	77.04	77.71	73.00	74.82	73.90
提案手法 ($\lambda = 1$)	78.37	77.00	77.68	72.56	74.36	73.45
人間	87.77	88.28	88.02	79.74	77.29	78.49

性がある。そこで、吉田ら [22] の手法を参考に、使用する漸進的なデータには 5% の制限を設けることとした。

性能の評価のため、以下の手法を比較手法として設定した。

提案手法 ($\lambda = 1$): 提案手法において、 $\lambda = 1$ とすることで主辞予測の学習を行わずに漸進的係り受け解析だけを学習させる手法。

Shift-Reduce (SR) 法: 日本語係り受け解析を行う Shift-Reduce アルゴリズム [11] を拡張して、漸進的係り受け解析と未入力の係り先文節の主辞予測を同時に行う手法 [5]。提案手法と同様に $\lambda = 1$ のときの結果も比較対象とする。

人間: 評価用データに対して人間 1 名が未入力文節予測と漸進的係り受け解析を図 1 の構造に基づいて行ったもの [26]。

漸進的係り受け解析の評価では、係り先が既入力の文節と未入力の文節に分類し、それぞれについて再現率、適合率、および F 値を計算した [6]。再現率は正解の係り受け構造のうちモデルが正しく解析したものの割合、適合率はモデルが解析した係り受け構造のうち正しいものの割合を示す。

未入力の係り先文節の主辞予測については、再現率、適合率、および F 値を用いて評価した [5]。再現率は、正解の係り受け構造における係り先が未入力である係り元文節のうち、係り先文節の主辞トークンを正しく予測できた係り元文節の割合である。適合率は、モデルが係り先を未入力と解析した係り元文節のうち、係り先文節の主辞トークンを正しく予測できた係り元文節の割合である。正解か否かの判定は、提案手法の出力の top- k ($k = 1, 3, 5$) に正解の主辞が含まれているか否かでを行った。

モデルの実装には Pytorch²⁾ を利用した。また、エンコーダには東北大学が公開している事前学習済み日本語 BERT モデル³⁾ を使用した。開発データを用

2) <https://pytorch.org/>

3) <https://huggingface.co/tohoku-nlp/bert-base-japanese-whole-word-masking>

表 2 未入力の係り先文節の主辞トークン予測結果

手法	再現率	適合率	F 値
Shift-Reduce 法 (top-1)	6.11	6.16	6.13
Shift-Reduce 法 (top-3)	14.72	14.85	14.78
Shift-Reduce 法 (top-5)	20.63	20.82	20.72
提案手法 (top-1)	7.80	7.99	7.90
提案手法 (top-3)	14.77	15.14	14.96
提案手法 (top-5)	18.78	19.25	19.01
人間 (top-1)	11.26	10.91	11.08

いてハイパーパラメータを調整した結果、バッチサイズを 8、学習率を $5e-5$ 、エポック数を 7、損失関数の重み係数 λ を 0.9 に設定した。評価では、乱数シードを変更し、各モデルをそれぞれ 5 つ用意し、各評価指標の平均値を測定した。

4.2 実験結果

4.2.1 漸進的係り受け解析の結果

表 1 に漸進的係り受け解析の評価結果を示す。提案手法と Shift-Reduce 法の F 値を比較すると、係り先が未入力の文節に対してはほぼ同等の性能を示した一方で、係り先が既入力の文節に対しては、提案手法の性能が低い結果となった。Shift-Reduce 法では、アルゴリズムの特性上、各文節が他の文節に係るか否かを 2 値分類として判定する。一方、提案手法は head selection 問題として定式化しており、入力されている文節の数が n 個の場合、係り先の候補は $2n - 1$ 個に増加する。このため、提案手法では係り先を正しく解析する難易度が高くなり、係り先が既入力の場合に性能が低下したと考えられる。

また、主辞予測を行わない設定である $\lambda = 1$ と比較した場合、提案手法は全ての指標で性能が向上しており、主辞予測が係り受け解析の性能向上に寄与していることが示された。

人間の性能との比較では、提案手法の精度との差は大きく、改善の余地があることが示された。

4.2.2 未入力の係り先文節の主辞予測の結果

表 2 に未入力の係り先文節の主辞予測の評価結果を示す。提案手法は、top-1 および top-3 において Shift-Reduce 法よりも高い精度を示し、それぞれ F 値が 1.77, 0.18 ポイント上回る結果となった。特に top-1 において、提案手法が Shift-Reduce 法を大きく上回る精度を示し、より正確に未入力の係り先文節の主辞を予測できることを確認した。Shift-Reduce 法が未入力の係り先文節の主辞予測を各未入力文節

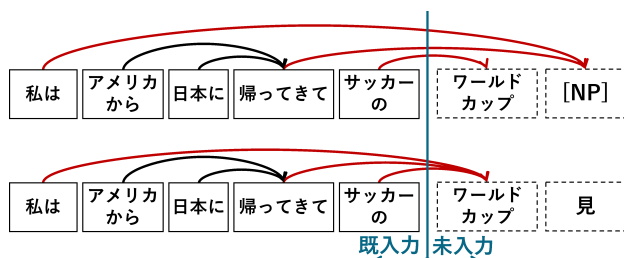


図3 漸進的係り受け解析と主辞予測が整合的でない例

ごとに独立して行い入力文節列の一部しか使用していないのに対し、提案手法では入力された文節情報を全て利用して各未入力文節の主辞トークンの予測を同時に行う。この仕組みにより、提案手法は、入力文節列全体の情報をより効果的に利用することができ、未入力の係り先文節の主辞予測において高い性能を発揮したと考えられる。

一方で、人間の性能と比較した場合、提案手法は未入力の係り先文節の主辞予測においても人間の性能には及んでいないことがわかる。このことから、未入力の係り先文節の主辞予測に関しても、提案手法の性能は改善の余地がある。

4.3 考察

提案手法は、漸進的係り受け解析と主辞予測を独立したタスクとして定式化し、マルチタスク学習を行うことで、それらを同時に実現する。しかし、2つのタスクの出力が必ずしも互いに整合的であるとは限らない。図3に示すように、未入力の係り先と解析された文節（図3上図では最右の点線四角の文節）に対して、主辞予測の結果、[NP]を予測する場合や、[NP]以外（図3下図では「見」）を予測した文節が係り先として選択されず、その文節を係り先とする係り元文節が存在しない場合がある。

そこで、本節では、提案手法における2つのタスクの出力がどの程度整合的に機能しているかを分析するために、以下の2つの観点から考察を行った。

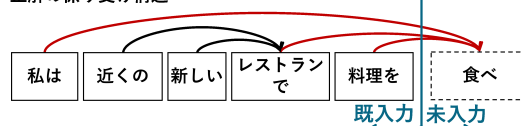
- モデルが未入力の係り先とした [UND] に対して、[NP] 以外を予測をしているか
- [NP] 以外を予測した [UND] に対して、その [UND] を係り先とする係り元文節がモデルの出力に存在するか

結果を表3に示す。この表から、提案手法は約90%の割合で未入力の係り先文節に対して [NP] 以外を予測していることがわかる。また、[NP] 以外を予測した [UND] に対して、その [UND] を係り先と

表3 漸進的係り受け解析と主辞予測が整合的であるかの分析結果

	割合 (%)
未入力の係り先文節への [NP] 以外の予測割合	89.57
[NP] 以外を予測した [UND] を係り先とする係り元文節が存在する割合	99.27

正解の係り受け構造



モデルが予測した係り受け構造

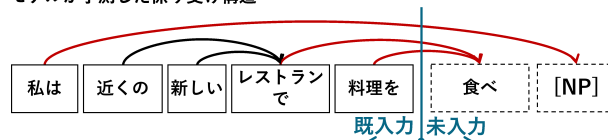


図4 正解の係り受け構造よりも多くの未入力の係り先文節があるとして解析した例

する係り元文節がモデルの出力に存在する割合は99%を超えている。これにより、提案手法は、未入力の係り先文節と解析した [UND] の多くに対して [NP] 以外を主辞として予測し、その [UND] を係り先とする未入力の係り先文節が存在すると解析していることを確認した。

一方で、未入力の係り先文節として解析された [UND] の約10%に対して、[NP] が予測された。例えば、図4のように、正解の係り受け構造よりも多くの未入力の係り先文節があると解析した場合、係り元文節が存在するにも関わらず、[NP] と予測された [UND] がモデルの出力に含まれることがあった。主辞予測の学習は正解の係り受け構造に基づいて行っている。すなわち、正解の係り受け構造における未入力の係り先文節の数を越えた [UND] は [NP] を出力するように学習されるため、図4のような現象が生じたものと考えられる。

5 おわりに

本論文では、漸進的係り受け解析と未入力の係り先文節の主辞予測を End-to-End で同時に行う手法を提案した。実験の結果、提案手法は未入力の係り先文節の主辞予測において、従来手法を上回る精度を達成した。一方で、漸進的係り受け解析において、係り先が既入力の文節に対する係り受け解析性能が従来手法よりも低下する結果が示された。

今後の課題として、係り受け解析と主辞予測をより密接に統合した学習方法の検討が挙げられる。

謝辞

本研究は JSPS 科研費 JP19K12127, JP24K15076 の助成を受けたものです。

参考文献

- [1] Dan Liu, Mengge Du, Xiaoxi Li, Ya Li, and Enhong Chen. Cross attention augmented transducer networks for simultaneous translation. In **Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)**, pp. 39–55, November 2021.
- [2] Tu Anh Nguyen, Eugene Kharitonov, Jade Copet, Yossi Adi, Wei-Ning Hsu, Ali Elkahky, Paden Tomasello, Robin Algayres, Benoît Sagot, Abdelrahman Mohamed, and Emmanuel Dupoux. Generative spoken dialogue language modeling. **Transactions of the Association for Computational Linguistics**, Vol. 11, pp. 250–266, 2023.
- [3] Stelios Piperidis, Iason Demiros, Prokopis Prokopidis, Peter Vanroose, Anja Hoethker, Walter Daelemans, Elsa Sklavounou, Manos Konstantinou, and Yannis Karavidas. Multimodal, multilingual resources in the subtitling process. In **Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)**, pp. 205–208, May 2004.
- [4] Tomohiro Ohno, Masaki Murata, and Shigeki Matsubara. Line-feed insertion into Japanese spoken monologue for captioning. In **Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing (ACL-IJCNLP 2009)**, pp. 531–539, August 2009.
- [5] 橋本優希, 大野誠寛, 松原茂樹. 日本語講演文に対する漸進的係り受け解析と文節主辞トークンの入力予測. 第 23 回情報科学技術フォーラム, pp. 333–336, 2024.
- [6] 橋本優希, 大野誠寛, 松原茂樹. 漸進的係り受け解析における BERT を用いた未入力文節との構文的関係の同定. 情報処理学会第 85 回全国大会講演論文集, No. 2, pp. 803–804, 2023.
- [7] Tomohiro Ohno and Shigeki Matsubara. Dependency structure for incremental parsing of Japanese and its application. In **Proceedings of the 13th International Conference on Parsing Technologies (IWPT 2013)**, pp. 91–97, November 2013.
- [8] 加藤芳秀, 松原茂樹, 外山勝彦, 稲垣康善. 主辞情報付き文脈自由文法に基づく漸進的な依存構造解析. 電子情報通信学会論文誌, Vol. J86-D-II, No. 1, pp. 84–97, January 2003.
- [9] Richard Johansson and Pierre Nugues. Incremental dependency parsing using online learning. In **Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2007)**, pp. 1134–1138, June 2007.
- [10] Joakim Nivre. Algorithms for deterministic incremental dependency parsing. **Computational Linguistics**, Vol. 34, No. 4, pp. 513–553, 2008.
- [11] 颯々野学. 日本語係り受け解析の線形時間アルゴリズム. 自然言語処理, Vol. 14, No. 1, pp. 3–18, 2007.
- [12] Xingxing Zhang, Jianpeng Cheng, and Mirella Lapata. Dependency parsing as head selection. In **Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EMNLP 2017)**, pp. 665–676, April 2017.
- [13] Zuchao Li, Jiaxun Cai, Shexia He, and Hai Zhao. Seq2seq dependency parsing. In **Proceedings of the 27th International Conference on Computational Linguistics (COLING 2018)**, pp. 3203–3214, August 2018.
- [14] Michalina Strzyz, David Vilares, and Carlos Gómez-Rodríguez. Viable dependency parsing as sequence labeling. In **Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)**, pp. 717–723, June 2019.
- [15] Ophélie Lacroix. Dependency parsing as sequence labeling with head-based encoding and multi-task learning. In **Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest 2019)**, pp. 136–143, August 2019.
- [16] Carlos Gómez-Rodríguez, Michalina Strzyz, and David Vilares. A unifying theory of transition-based and sequence labeling parsing. In **Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020)**, pp. 3776–3793, December 2020.
- [17] Nikita Kitaev and Dan Klein. Tetra-tagging: Word-synchronous parsing with linear-time inference. In **Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)**, pp. 6255–6261, July 2020.
- [18] Afra Amini and Ryan Cotterell. On parsing as tagging. In **Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022)**, pp. 8884–8900, December 2022.
- [19] Carlos Gómez-Rodríguez, Diego Roca, and David Vilares. 4 and 7-bit labeling for projective and non-projective dependency trees. In **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)**, pp. 6375–6384, December 2023.
- [20] Afra Amini, Tianyu Liu, and Ryan Cotterell. Hexatagging: Projective dependency parsing as tagging. In **Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (ACL 2023)**, pp. 1453–1464, July 2023.
- [21] 柴田知秀, 河原大輔, 黒橋禎夫. BERT による日本語構文解析の精度向上. 言語処理学会 第 25 回年次大会 発表論文集, pp. 205–208, 2019.
- [22] 吉田あいり, 河原大輔. 構造的曖昧性に基づく読みづらさの検出. 言語処理学会 第 28 回年次大会 発表論文集, pp. 425–429, 2022.
- [23] Kiyotaka Uchimoto, Satoshi Sekine, and Hitoshi Isahara. Japanese dependency structure analysis based on maximum entropy models. In **Proceedings of the 9th Conference of the European Chapter of the Association for Computational Linguistics (EACL 1999)**, pp. 196–203, 1999.
- [24] Shigeki Matsubara, Akira Takagi, Nobuo Kawaguchi, and Yasuyoshi Inagaki. Bilingual spoken monologue corpus for simultaneous machine interpretation research. In **Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)**, pp. 153–159, 2002.
- [25] 大野誠寛, 松原茂樹. 文節間の依存・非依存を同定する漸進的係り受け解析. 電子情報通信学会論文誌, Vol. J98-D, No. 4, pp. 709–718, 2015.
- [26] Hiroki Unno, Tomohiro Ohno, Koichiro Ito, and Shigeki Matsubara. Human performance in incremental dependency parsing: Dependency structure annotations and their analyses. In **Proceedings of the 38th Pacific Asia Conference on Language, Information and Computation (PACLIC 2024)**, December 2024.