

作業記憶の発達的特性が言語獲得の臨界期を形成する

三田雅人^{1,2} 吉田遼¹ 深津聡世¹ 大関洋平¹

¹ 東京大学 ² 株式会社サイバーエージェント

mita_masato@cyberagent.co.jp

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

概要

大規模言語モデル (LLM) は汎用的な言語能力を持つが、言語獲得効率では人間と大きな乖離がある。本研究では、人間が特定の時期において言語獲得が特に効率的に進むとされる臨界期における作業記憶の発達的特性を言語モデルに組み込む手法を提案した。提案手法では、学習初期に“作業記憶”を制限し、学習が進むにつれて指数関数的に制限を緩和する仕組みを導入する。文法評価ベンチマークで性能を評価した結果、提案手法は従来手法を上回る性能を示し、特に、局所的な規則性と文全体の複雑な関係を含む文法項目で顕著な改善が見られた。これらの知見は、データ効率の高い LLM 設計の新たな指針を提供するだけでなく、人間の言語獲得における臨界期仮説を支持する重要な間接証拠となる。

1 はじめに

大規模言語モデル (LLM) は、あたかも人間のような汎用的な言語能力を有する一方、言語獲得効率においては依然として人間には遠く及ばない。LLM は多くの評価指標において人間と同等の性能を達するために人間と比較して 3~4 桁多いデータ量を必要とすることが指摘されている [1]。この LLM と人間のデータ効率性の乖離は、LLM がスケールアップに依存している現状を反映しており、学習効率を抜本的に改善する潜在的な余地があることを示唆していると同時に、人間の言語処理や獲得過程からの「教訓」が存在しうることを示唆している。

人間の言語獲得効率についての重要な理論として、臨界期仮説 [2] がある。臨界期仮説は、言語を効率的に習得できる特定の時期が存在し、この時期を過ぎるとその能力が低下するという理論である。幼少期に言語を聞く機会が制限された事例や第二言語習得における年齢の影響など、多くの研究が臨界期の存在を支持している [3]。しかし、幼児が大人

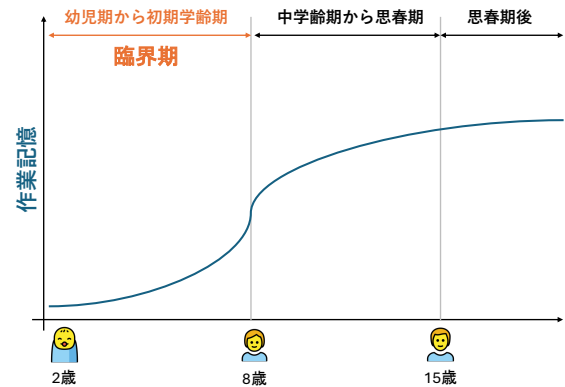


図1 作業記憶の発達過程の概要図

より効率よく言語を習得できる理由は未解明な部分が多い。言語獲得における臨界期の有力な説明の一つとして、幼児の認知的な制約 (例: 短期記憶の容量や注意の範囲) がむしろ言語学習に有利に働くとされる Newport の Less is more 仮説 [4] がある。この仮説では、幼児は処理可能な情報が限られるため、言語の基本的なパターンや構造 (例: 文法規則) を効率的に抽出できる一方、大人は認知能力が高いがゆえに複雑な情報に気を取られ、規則の学習が妨げられると説明される。

本研究では、Less is more 仮説から着想を得て、人間の認知発達的特性、特に作業記憶の発達的特性を言語モデルに組み込むことで効率的な言語獲得が誘発されるかを調査する。具体的には、臨界期に相当する時期の作業記憶の指数関数的増加を模倣した言語モデルを設計し、これが学習効率にどのような影響を与えるかを分析する。子ども向け発話 (Child-Directed Speech; CDS) データセット [5] を用いて訓練した GPT-2 [6] をもとに、CDS に特化した文法評価ベンチマークである Zorro [7] を用いて評価実験を行ったところ、学習初期から一定の作業記憶を有するモデルに比べ、学習初期は作業記憶を制限し、学習が進むにつれて指数関数的に制限を緩和する認知発達的に妥当なモデルの方が性能が高い結

果となった。これらの知見はデータ効率の高い言語モデル設計に向けた新たな指針を提供するとともに（自然言語処理分野への貢献）、長年議論が続いている人間の言語獲得における臨界期仮説に対する間接証拠を提供するものである（認知科学分野への貢献）。

2 関連研究

2.1 言語獲得における臨界期仮説

臨界期仮説は、言語獲得が特定の時期において特に効率的に進む一方、この期間を過ぎるとその能力が低下するという理論である。心理言語学分野では、第二言語習得における発音や文法の習熟が思春期以降に急激に低下することが指摘されており、こうした現象は生物学的な臨界期の存在を支持する証拠とされている [8, 9]。また、神経科学分野では、脳の可塑性が発達とともに減少することや神経伝達物質の変化が臨界期の終了に寄与する可能性が示唆されている [2]。この仮説を支持する生得主義的立場では、言語獲得は遺伝的にプログラムされた能力であり、臨界期はその生物学的基盤とされる [10, 11]。経験主義的立場からは、臨界期の効果は社会的・環境的要因、特に言語的インプットの量と質によっても説明可能であるとされる [12, 13]。しかし、臨界期が言語能力に与える影響の範囲や終了時期、生物学的メカニズムと環境要因の相対的な寄与については依然として議論が続いている。例えば、言語獲得の臨界期の終了時期については、Lenneberg [2] は脳の側性化が思春期までに完了し、これが言語獲得能力の低下と関連することから臨界期は思春期（約 12～13 歳）まで続くと主張している。一方で、Krashen [14] は、この Lenneberg の仮説を再検討し、臨界期の基盤となる脳の側性化の発達が約 5 歳までに完了することを示し、臨界期が思春期よりも早い時期に終了する可能性を指摘している。

2.2 獲得理論における言語モデルの役割

近年、計算モデルは言語獲得のメカニズム解明において重要な役割を果たしている。これらのモデルは、人間被験者では難しい学習メカニズムや環境の制御を可能にし、「刺激の貧困」論のような理論的主張を検証するために用いられている [15]。例えば、McCoy ら [16]、Wilcox ら [17]、Warstadt ら [1] は、言語モデルを用いて言語獲得に関する仮説を直接

検証し、モデルが学習可能性に関する存在証明を提供できることを示している。このような研究は、計算モデルを通じて言語獲得の理論的議論を深める試みとして注目されており、言語獲得の臨界期も例外ではない。Constantinescu ら [18] は、特に、第二言語 (L2) の習得¹⁾と第一言語 (L1) の喪失²⁾における臨界期現象に焦点を当てた実験を行い、言語モデルが人間の学習者と異なり、臨界期効果を自然に示さないことを明らかにした。また、破滅的忘却の軽減に一般的に用いられる正則化手法 Elastic Weight Consolidation [19] を導入することで、人間の脳の可塑性を模倣し、臨界期効果を人工的に再現可能であることを示した。これらの結果は、臨界期効果が統計的学習の必然的な結果ではなく、生得的な要因による可能性を支持するものである。Constantinescu らの研究では L2 習得と L1 喪失における臨界期効果を再現することを目的にしているのに対して、本研究では L1 獲得における臨界期効果が再現することを目的としている点が大きく異なる。

3 発達的に妥当な作業記憶

3.1 作業記憶の発達的变化

人間の作業記憶は発達段階に応じて顕著な変化を示し、概ね以下の 3 つの時期に分類される：

- **幼児期から初期学齢期 (2～7 歳)** この時期には、情報の保持能力と処理能力が大きく向上する。特に、認知資源の拡張が観察される [20, 21]。
- **中学齢期から思春期 (8～14 歳)** 作業記憶の発達は引き続き進行するが、脳の成熟に伴い成長速度は緩やかになる [22, 21]。
- **思春期後 (15 歳以上)** この段階では、作業記憶の能力が成人レベルに達し、発達はほぼ停滞する [23, 22]。

これらを踏まえると人間の作業記憶は図 1 のような軌跡を辿ると考えられる。ここで、2.1 節で述べた通り、臨界期の時期については議論が分かれているが、(1) 両者の主張が最低限合意できること、および (2) 作業記憶の発達段階との対応がとれることを踏まえて本研究では臨界期の終了時期をおおよそ 7 歳程度までと仮定する。このとき、臨界期現象が

1) L2 への接触開始年齢が遅いほど学習が難しくなり、熟達度が低下する現象

2) L1 への接触が途絶える年齢が早いほど、L1 を忘れる（喪失する）可能性が高くなる現象

起こる時期の人間の作業記憶は線形増加というよりは指数関数的増加であると既定できる。

3.2 作業記憶の制限

作業記憶の制限に関するモデリングに、Attention with Linear Biases (ALiBi) [24] を導入する。ALiBi は、Transformer モデルにおいて位置埋め込みを用いず、注意スコアに対して距離に応じた線形ペナルティを追加する手法である。この方法では、クエリとキーのドット積に固定のバイアスを加える。具体的には、入力長を L とした場合の注意スコア (Attention Score) は次式で計算される：

$$\text{Attention Score} = \text{softmax}(q_i K^T + m \cdot B),$$
$$B = \begin{bmatrix} -(i-1) & -(i-2) & \dots & 0 \end{bmatrix}. \quad (1)$$

ここで $q_i \in \mathbb{R}^{1 \times d}$ はクエリ、 $K \in \mathbb{R}^{L \times d}$ はキー、 m はヘッド固有のスカラー係数 (スロープ)、 B はクエリとキーの相対距離に基づくバイアス行列である。スロープ m の値は幾何数列に基づいて設定され、例えば、8 ヘッドの場合には m の値は $1, \frac{1}{2}, \frac{1}{4}, \dots, \frac{1}{128}$ のような等比数列となる。

ALiBi では距離の大きいクエリとキーのペアに対する注意スコアが抑制されるため、モデルに「近近性」のバイアスを導入できる。ALiBi は元々は Transformer モデルの外挿能力を向上させることを目的に提案されたが、モデル訓練時の注意スコア計算に使用することで人間の読み時間に近い傾向のサブライザルを推定することが可能になることも報告 [25] されており、人間らしい記憶の減衰 (制限) に関するモデリングが期待できる。

しかし、ALiBi におけるスロープ m は各注意ヘッドにおいて固定のため、このままでは年齢とともに作業記憶が増加 (記憶の減衰幅が低下) するという発達の特性を十分に反映できない。そのため、本研究では ALiBi におけるスロープ m をエポックの進行に伴い指数関数的に減少させることで、臨界期が形成されるとされる人間の幼児期から初期学齢期の発達過程における作業記憶の指数関数的増加を再現する手法 (DYNAMICLIMIT-EXP) を提案する。この手法では、エポック数 t に基づいて式 (1) のスロープ m を次式のように更新する：

$$m_t = m_0 \cdot t^{-\lambda}, \quad (2)$$

ここで、 m_0 は初期スロープ、 $\lambda > 0$ は減少率を制

御するハイパーパラメータ、 t は現在のエポック数を表す。これにより、モデルは初期段階では近距離の注意を重視し、学習が進むにつれて長距離の依存関係に注意を向けられるようになる。

4 実験

人間の作業記憶の発達の特性をフルスクラッチで学習した言語モデルに組み込むことで効率的な言語獲得が促進されるか、すなわち L1 獲得における臨界期効果が再現されるかについて実験する。

4.1 モデル

実験に使用する言語モデルとして、本研究では GPT-2 [6] を採用する。RoBERTa に基づくモデルを使用する既存研究 [7, 1] も存在するが、(1) 人間の短期記憶の制限 (減衰) のモデリングには双方向ではなく単方向 (左から右) での予測動作が妥当であること、(2) 現在主流の LLM は GPT に基づくアーキテクチャが採用されていることの大きく 2 点が選定理由である。訓練データセットには、CDS データセットの一つである AO-CHILDES [5]³⁾ を用いる。AO-CHILDES は、英語圏のおよそ 1 歳から 6 歳までの子ども向け発話が年齢順に収録された 5M 単語規模のデータセットであり、子どもの年齢層の違いや話者の違い、場面の違いなどの外的要因の統制が図られている。GPT-2 モデルの詳細な学習設定は付録 A に記載する。

本実験では、提案手法である DYNAMICLIMIT-EXP に加えて、以下の 3 つのベースラインモデルを用意し、作業記憶における制限方法の違いによる学習効果を精緻に分析する：

- **NOLIMIT**: 記憶の制限をかけないモデル。学習初期から作業記憶が一定であり、思春期以降に観察される発達が成熟した作業記憶を模倣している。
- **STATICLIMIT**: 注意スコア計算時に通常の ALiBi を適用したモデル。学習初期から学習後期にかけて一定の記憶制限がかけられている。
- **DYNAMICLIMIT-LINEAR**: ALiBi のスロープ m を学習の進行に伴い線形に減少させることで、中学齢期から思春期の期間に観察される作業記憶の線形増加を模倣するモデル。

3) <https://github.com/UIUCLearningLanguageLab/AOCHILDES>

表1 Zorroにおける各言語モデルの精度 (%) .OVERALL は各文法項目のスコアのマクロ平均を表す。

Model	OVERALL	D-N AGR	S-V AGR	ANA. AGR	ARG. STR	BINDING	CASE	ELLIPSIS	FILLER. GAP	IRREGULAR	ISLAND	LOCAL. ATR	QUANTIFIERS	NPI
NoLIMIT	52.0	48.7	49.9	50.9	30.0	63.6	15.9	69.5	83.0	46.3	54.1	64.6	31.7	80.5
STATICLIMIT	47.3	49.8	49.5	52.9	16.9	42.2	6.7	72.3	82.5	52.6	57.6	64.4	34.1	50.0
DYNAMICLIMIT-LINEAR	46.7	50.3	50.6	48.7	39.2	48.8	28.7	25.8	64.4	51.1	50.1	68.6	29.3	46.8
DYNAMICLIMIT-EXP	56.5	52.1	50.2	47.6	49.1	59.2	56.8	67.5	92.9	49.9	60.6	50.1	50.0	55.2

4.2 評価

モデルの文法能力を測定するデータセットとして Zorro [7] を用いる。Zorro は英語の文法項目ごとに容認可能な文と不可能な文からなるミニマルペアを用いて文法能力を評価する BLiMP [26] に触発されて構築された文法評価ベンチマークである。BLiMP に比べて使用語彙や文法項目をより CDS に適応させたものであり、本研究では中分類（以降は文法項目と記す）ごとのスコアとそれらのマクロ平均を報告する。以下は「主語と動詞の一致 (Subject-verb agreement)」という項目のペアの例である⁴⁾。

- (1) a. The **lie** on the foot is flat.
 b.*The **lies** on the foot is flat.

容認可能な文と不可能な文をそれぞれモデルに入力し、前者により高い確率が付与されたペアの割合を計算することで文法性判断スコアを得る。

4.3 結果

表1 は Zorro における各言語モデルの精度を示している。各文法項目におけるマクロ平均 OVERALL の結果から、DYNAMICLIMIT-EXP は、NoLIMIT より 4.5 ポイント性能が向上し、作業記憶を初期に制限し、指数関数的に緩和する設計が効果的であることがわかった。この結果は、言語モデルが人間と同様に L1 獲得の臨界期効果を再現したことを示唆する。一方、STATICLIMIT や DYNLIMIT-LINEAR は NoLIMIT より性能が低下しており、単に記憶制限を設けるだけでは効果がなく、指数関数的緩和が重要であることを示している。

ここで、なぜ線形増加ではなく指数関数的増加が効果的であるのかという疑問が残る。NoLIMIT から DYNAMICLIMIT-LINEAR への精度の上昇幅と比較して、DYNAMICLIMIT-EXP による精度の大幅な向上が顕著な文法項目として、D-N AGR, ARG. STR, CASE, FILLER. GAP, ISLAND, QUANTIFIERS が挙げられる。こ

4) Zorro に含まれる全文法項目を付録 B に示す

れらの項目の多くは、局所的な情報に基づく規則性と、文全体の構造や文脈に基づく複雑な関係性の双方を内包する。例えば、ISLAND (島条件) は、特定の補文構造では要素の移動が可能だが、他の構造では不可能になる制約を示す：

- (2) a. Who did you say [that Mary saw _]?
 b.*Who did you say [whether Mary saw _]?

a の例は、「saw」と「Who」の局所的な依存関係を捉え、補文「that Mary saw _」が主節に適切に埋め込まれているため、文法的に正しい。一方、b の例では、「whether 節」が島条件を形成し、「Who」を移動させることは非文法的である。このような構造を処理するには、まず「saw」と「Who」の局所的な依存関係を解析し、次に補文の種類や主節との関係性を評価してその文法的な違いを理解する必要がある。これらのプロセスを直列的に処理することで、文の容認性を判断できる。ここで、DYNAMICLIMIT-EXP は、DYNAMICLIMIT-LINEAR と比較し、学習初期には記憶スパンが小さく、学習後期には急激に増加する特徴を持つ。この段階的な緩和は、初期に基本的なパターン抽出を優先し、後に複雑な規則をブートストラップ的に学習させることで規則の汎化を促進していると考えられる。

5 おわりに

本研究では、人間の作業記憶の発達特性を模倣した言語モデルを提案し、その有効性を示した。具体的には、Less is more 仮説に基づき、学習初期に作業記憶を制限し、進行に伴い指数関数的に緩和する手法を導入した。文法評価実験において、提案手法は記憶の制限をかけないベースラインを超える文法性能を示し、特に、局所的な情報に基づく規則性と、文全体の構造や文脈に基づく複雑な関係性の双方を内包する文法項目において顕著な改善が見られた。この結果は、作業記憶の発達が臨界期形成に重要である可能性を示唆し、データ効率の向上や言語獲得の理解に新たな知見を提供する。

謝辞

本研究は、JST さきがけ JPMJPR21C2 および JSPS 科研費 JP24H00087 の支援を受けたものです。

参考文献

- [1] Alex Warstadt, Aaron Mueller, Leshem Choshen, Ethan Wilcox, Chengxu Zhuang, Juan Ciro, Rafael Mosquera, Bhargavi Paranjabe, Adina Williams, Tal Linzen, and Ryan Cotterell. Findings of the BabyLM challenge: Sample-efficient pretraining on developmentally plausible corpora. In **Proceedings of the BabyLM Challenge at the 27th Conference on Computational Natural Language Learning**, pp. 1–34. Association for Computational Linguistics, December 2023.
- [2] E.H. Lenneberg. **Biological Foundations of Language**. Wiley, 1967.
- [3] Jacqueline S Johnson and Elissa L Newport. Critical period effects in second language learning: The influence of maturational state on the acquisition of english as a second language. **Cognitive Psychology**, Vol. 21, No. 1, pp. 60–99, 1989.
- [4] Elissa L. Newport. Maturational constraints on language learning. **Cognitive Science**, Vol. 14, No. 1, 1990.
- [5] Philip A. Huebner and Jon A. Willits. **Using lexical context to discover the noun category: Younger children have it easier**, pp. 279–331. *Psychology of Learning and Motivation - Advances in Research and Theory*. Academic Press Inc., January 2021.
- [6] Alec Radford, Jeff Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. 2019.
- [7] Philip A. Huebner, Elinor Sulem, Fisher Cynthia, and Dan Roth. BabyBERTa: Learning more grammar with small-scale child-directed language. In **Proceedings of the 25th Conference on Computational Natural Language Learning**, pp. 624–646. Association for Computational Linguistics, November 2021.
- [8] Sonia Tahta, Margaret Wood, and Kate Loewenthal. Foreign accents: Factors relating to transfer of accent from the first language to a second language. **Language and Speech**, Vol. 24, No. 3, pp. 265–272, 1981.
- [9] Mark S. Patkowski. The sensitive period for the acquisition of syntax in a second language. **Language Learning**, Vol. 30, No. 2, pp. 449–468, 1980.
- [10] Noam Chomsky. **Aspects of the Theory of Syntax**. The MIT Press, Cambridge, 1965.
- [11] Steven Pinker. **The Language Instinct: How the Mind Creates Language**. William Morrow and Company, 1994.
- [12] Jeffrey L. Elman, Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. **Rethinking Innateness: A Connectionist Perspective on Development**. MIT Press, 1996.
- [13] Mark Seidenberg and David Plaut. 2 progress in understanding word reading: Data fitting versus theory building. **From Inkmarks to Ideas: Current Issues in Lexical Processing**, 01 2006.
- [14] Stephen D. Krashen. Lateralization, language learning, and the critical period: Some new evidence. **Language Learning**, Vol. 23, No. 1, pp. 63–74, 1973.
- [15] Alexander Clark and Shalom Lappin. **Linguistic Nativism and the Poverty of the Stimulus**. Wiley-Blackwell, 2011.
- [16] R. Thomas McCoy, Robert Frank, and Tal Linzen. Does syntax need to grow on trees? sources of hierarchical inductive bias in sequence-to-sequence networks. **Transactions of the Association for Computational Linguistics**, Vol. 8, pp. 125–140, 01 2020.
- [17] Ethan Gottlieb Wilcox, Richard Futrell, and Roger Levy. Using computational models to test syntactic learnability. **Linguistic Inquiry**, Vol. 55, No. 4, pp. 805–848, 10 2024.
- [18] Ionut Constantinescu, Tiago Pimentel, Ryan Cotterell, and Alex Warstadt. Investigating critical period effects in language acquisition through neural language models, 2024.
- [19] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. **Proceedings of the National Academy of Sciences**, Vol. 114, No. 13, pp. 3521–3526, 2017.
- [20] Nelson Cowan, Lara Nugent, Emily M. Elliott, Igor Ponomarev, and John Scott Sauls. The role of attention in the development of short-term memory: age differences in the verbal span of apprehension. **Child development**, Vol. 70 5, pp. 1082–97, 1999.
- [21] S. E. Gathercole, S. J. Pickering, B. Ambridge, and H. Wearing. The structure of working memory from 4 to 15 years of age. **Developmental psychology**, Vol. 40, No. 2, pp. 177–190, 2004. Gathercole, Susan E Pickering, Susan J Ambridge, Benjamin Wearing, Hannah 2004/2/26.
- [22] Beatriz Luna, Krista E. Garver, Trinity A. Urban, Nicole A. Lazar, and John A. Sweeney. Maturation of cognitive processes from late childhood to adulthood. **Child Development**, Vol. 75, No. 5, pp. 1357–1372, 2004.
- [23] Elizabeth R. Sowell, Doris A. Trauner, Anthony Collins Gamst, and Terry L. Jernigan. Development of cortical and subcortical brain structures in childhood and adolescence: a structural mri study. **Developmental Medicine & Child Neurology**, Vol. 44, , 2002.
- [24] Ofir Press, Noah Smith, and Mike Lewis. Train short, test long: Attention with linear biases enables input length extrapolation. In **International Conference on Learning Representations**, 2022.
- [25] Christian Clark, Byung-Doh Oh, and William Schuler. Linear recency bias during training improves transformers’ fit to reading times, 2024.
- [26] 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.

A ベースラインモデル実験設定

表2 GPT-2 モデルの学習設定 (ハイパーパラメータ) .

ハイパーパラメータ	値
モデル構造	GPT-2 (カスタム設定)
レイヤー数	4
アテンションヘッド数	4
埋め込み次元	256
ドロップアウト率	0.1
学習率 (η)	5×10^{-6}
重み減衰	0.01
バッチサイズ	512
勾配蓄積ステップ数	2
総エポック数	20
最大文長	32
学習率スケジューラー	コサイン (リスタートあり)
ウォームアップステップ数	総ステップ数の 10%
最適化器	AdamW, $\beta = (0.9, 0.999)$ and $\epsilon = 1e - 08$
トークナイザー	カスタム (AO-CHILDES で学習)
早期終了の許容エポック数	1 エポック
評価指標	Perplexity

B Zorro における文法項目詳細

表3 Zorro における各文法項目の説明. 各事例は元論文 [7] の Table5 に記載のものである.

中分類	小分類	容認可能な文	容認不可能な文
D-N AGR	noun-across_l_adjective noun-between_neighbors	<i>look at this purple thing .</i> <i>this color must be white .</i>	<i>look at this purple things .</i> <i>this colors must be white .</i>
S-V AGR	verb-across_prepositional_phrase verb-across_relative_clause verb-in_question_with_aux verb-in_simple_question	<i>the lie on the foot is flat .</i> <i>the book that i like is poor .</i> <i>where does the horse go ?</i> <i>where is the way ?</i>	<i>the lies on the foot is flat .</i> <i>the books that i like is poor .</i> <i>where does the horses go ?</i> <i>where is the ways ?</i>
ANA.AGR	pronoun_gender	<i>will Mark want himself ?</i>	<i>will Mark want herself ?</i>
ARG.STR	dropped_argument swapped_arguments transitive	<i>give me the poor boat .</i> <i>he made the slave her label .</i> <i>Philip thinks .</i>	<i>the poor boat gives me .</i> <i>the slave made her label he .</i> <i>Philip affected .</i>
BINDING	principle_a	<i>Ben thinks about himself calling this fuel .</i>	<i>Ben thinks about himself called this fuel .</i>
CASE	subjective_pronoun	<i>i brought the wolf my hill .</i>	<i>the wolf brought i my hill .</i>
ELLIPSIS	n_bar	<i>Mark fixed one worn canal and Roger fixed more .</i>	<i>Mark fixed one canal and Roger fixed more worn .</i>
FILLER.GAP	wh_question_object wh_question_subject	<i>Laura married the dinner that the wolf could close .</i> <i>Laura ended the finger that can make boats .</i>	<i>Laura married what the dinner could close the wolf .</i> <i>Laura ended who the finger can make boats .</i>
IRREGULAR	verb	<i>Michael chose the good one some time ago .</i>	<i>Michael chosen the good one some time ago .</i>
ISLAND	adjunct_island coordinate_structure_constraint	<i>who should William have without watching the baby ?</i> <i>who must Philip and the dinosaur turn ?</i>	<i>who should William have the baby without watching ?</i> <i>who must Philip turn and the dinosaur ?</i>
LOCAL.ATR	in_question_with_aux	<i>is the whale getting the person ?</i>	<i>is the whale gets the person ?</i>
NPI	matrix_question only_npi_licensor	<i>does her boat ever play with the growth ?</i> <i>only Mark ever finds some suit .</i>	<i>her boat does ever play with the growth ?</i> <i>even Mark ever finds some suit .</i>
QUANTIFIERS	existential_there superlative	<i>there are many books about soft birds .</i> <i>no pig could stand on top of more than six days .</i>	<i>there are most books about soft birds .</i> <i>no pig could stand on top of at least six days .</i>