

# 大規模言語モデルの衝撃を生き残る統語理論を考える

磯野真之介<sup>1</sup> 梶川康平<sup>2</sup><sup>1</sup> 国立国語研究所 <sup>2</sup> ジョージタウン大学

s-isono@ninjal.ac.jp kk1571@georgetown.edu

## 概要

離散記号の再帰的な組み合わせとしての統語構造は、大規模言語モデル (LLM) 全盛の今日にあっても、LLM や人間の脳における言語の実装を分析する起点として、重要な役割を果たしうると指摘されている。ところが、従来の統語理論の「主流」たる Chomsky 派生成文法は、多方面から厳しく批判されている。この現状を踏まえ、本稿は、LLM 以降の統語理論として、単語の組み合わせに対する制約を型論理とみなす枠組みが有効であると主張する。型論理に基づく統語理論は、従来の統語研究がもたらした知見を生かし、Chomsky らが重視してきた簡潔性の要求に応えつつ、今後の統語理論に求められる形式的厳密性、逐次的処理との親和性、進化や獲得の問題との接続可能性、大規模データによる検証可能性といった諸特性を備えた枠組みである。

## 1 はじめに

大規模言語モデル (LLM) の成功は、言語の科学にも衝撃を与えた。その衝撃を端的に表し、話題を呼んだのが Piantadosi による「現代の言語モデルは Chomsky の言語へのアプローチを反駁する」<sup>1)</sup>と題された論考 [1] である。LLM は真正な言語の理論であり、従来の理論言語学 (特に統語論) で主流とされてきた Chomsky の枠組みは、「核心的な主張から特定の洞察や原理、構造、過程」(要旨) に至るまで覆された、と宣告したのである。この強烈な主張には当然反論もあり、現在まで議論が続いている。Futrell と Mahowald [2] は、LLM と従来の言語学を対置する見方を批判し、従来の言語学が見出してきた言語的構造は、言語現象の予測や記述の圧縮を可能にする抽象化という意味で実在するパターン [3] であり、LLM とともに重要であり続けると指摘する。

しかし、「主流」生成文法とも呼ばれる Chomsky の統語理論 [4, 5] は、脳や LLM の分析の起点とし

1) 訳は筆者、以下同じ。

での役割を担うには問題が多い。<sup>2)</sup> Piantadosi の批判を待つまでもなく、形式性の欠如 [6] や逐次的処理との接続の軽視 [7]、それに由来する、大規模データを用いた検証といった計算言語学的パラダイムとの相性の悪さ [8] などは言語学者からも批判されてきた。それでも、Chomsky 派が主流とされるがゆえに、Piantadosi 自身は批判の対象を「Chomsky のアプローチ」と定めているにも関わらず、その論考に触発された論争は「言語モデルと言語学を対置する」([2], p. 29) 雰囲気を持ち、「(生成) 言語学の終わり」([8], タイトル) さえ議論されている。

主流生成文法が仮に終わりだとしても、理論言語学が明らかにしてきた統語構造の有効性は、少なくとも実在のパターンとして残るはずである。この問題意識のもと、本稿は、**統語をある種の型論理とみなす範疇文法**の枠組みが、LLM 以降の時代の統語理論として有効であると主張する (cf.[9])。2 節は型論理範疇文法を導入する。3 節は、同文法が今後の統語理論に求められる諸特性を備えていることを述べ、今後の課題にも触れる。

## 2 型論理としての統語

範疇文法は Ajdukiewicz [10] が文法と算術との間に並行性を見出したことに始まる。Lambek [11, 12] が論理として再定式化し、そこから発展した文法群は特に**型論理文法**と呼ばれる。よく知られる組合せ範疇文法 (CCG) [13]<sup>3)</sup> は、型論理文法と対置されることもあるが、論理としての定式化も可能である [14]。本稿では、Lambek 流の型論理文法を中心に扱うが、CCG も逐次処理などの観点で重要であり、両者の関係にも触れる (3.5 節)<sup>4)</sup>。

型論理文法では、言語普遍的な統語計算の仕組み

2) あくまで枠組みの問題であって、主流の枠組みを用いつつも、具体的な現象の分析において言語学的に重要な洞察をもたらしている論文はいくらでもある。

3) 日本語での導入としては [15, 16] がある。

4) 以下では、範疇文法一般に当てはまる記述についても、混乱を避けるために「型論理文法」に統一する。

として何らかの形式論理を仮定する。この論理における式は言語表現の型とみなされる。個別言語の文法は、原始式の集合および語彙への式の割当を与えることで、論理と現実の言語表現を接続する。

型論理文法の出発点となった Lambek の論理体系  $L^5$  [11] では、原始式の集合  $P$  が与えられたとき、式の集合  $F$  は次のように定められる：

$$F ::= P \mid F \setminus F \mid F / F \mid F \bullet F$$

$b \setminus a$  と  $a / b$  はそれぞれ型  $b$  の項を左や右にとって型  $a$  の表現となる表現の型<sup>6)</sup>、 $a \bullet b$  は型  $a$  と型  $b$  の表現を並べた表現の型である。記法は先述の算術との並行関係に由来する。すなわち型  $b \setminus a$  および  $a / b$  は分数  $\frac{a}{b}$  に対応し、 $a \bullet b$  は積に対応する。命題論理でいえばそれぞれ含意と連言、プログラミングでいえば関数型とペア型である。ただし  $L$  では  $\bullet$  は結合的・非可換である。

論理的推論はシーケント  $\Gamma \vdash a$  についてなされる。  $\Gamma$  はカンマで区切られた式の列である。言語学的には、 $\Gamma \vdash a$  は「 $\Gamma$  内の各型の表現を並べると型  $a$  の表現になる」を意味する。証明体系は等価なもの複数あるが、Gentzen 流（各結合子をシーケントの左と右に導入する規則がそれぞれある）の  $L$  の公理と推論規則は以下の通りである：<sup>7)</sup>

$$\frac{}{a \vdash a} [\text{ax}]$$

$$\frac{\Gamma[a] \vdash c \quad \Delta \vdash b}{\Gamma[\Delta, b \setminus a] \vdash c} [\setminus L] \quad \frac{b, \Gamma \vdash a}{\Gamma \vdash b \setminus a} [\setminus R]$$

$$\frac{\Gamma[a] \vdash c \quad \Delta \vdash b}{\Gamma[a / b, \Delta] \vdash c} [/L] \quad \frac{\Gamma, b \vdash a}{\Gamma \vdash a / b} [/R]$$

$$\frac{\Gamma[a, b] \vdash c}{\Gamma[a \bullet b] \vdash c} [\bullet L] \quad \frac{\Gamma \vdash a \quad \Delta \vdash b}{\Gamma, \Delta \vdash a \bullet b} [\bullet R]$$

具体例として英語の断片を見る。原始式として文  $s$ 、名詞句  $n$ 、普通名詞  $cn$  をおく。このとき次の証明は妥当な証明である。そして、最下段に付与された語彙を各式の一解釈として定めることで、「*Pat saw Kim* は文である」と言えることになる。

$$\frac{\frac{\frac{s \vdash s}{s \vdash s} [\text{ax}] \quad \frac{n \vdash n}{n \vdash n} [\text{ax}]}{n, n \setminus s \vdash s} [\setminus L] \quad \frac{}{n \vdash n} [\text{ax}]}{\frac{n, (n \setminus s) / n, n \vdash s}{\underbrace{\quad}_{\text{Pat}} \quad \underbrace{\quad}_{\text{saw}} \quad \underbrace{\quad}_{\text{Kim}}} [/L]}$$

型論理文法の特長の一つは、 $[\setminus R]$ 、 $[/R]$  規則による仮定的推論により、高階関数型が柔軟に使えること

5) 本稿の  $L$  の定式化は [17] に基づく。  
6) CCG では返り値が常に左に置かれる記法  $(a / b, a \setminus b)$  をとる。不幸な不一致だが、分数に由来する記法であることを踏まえ、返り値が常に上に置かれる記法をとる。  
7) カット規則は省略する。

にある。高階関数型は量子子や疑問文などの分析に有効である。例えば、次は  $L$  による英語の目的語関係節の分析である [18]：

$$\frac{\frac{\frac{n / cn, cn, cn \setminus cn \vdash n}{n / cn, cn, (cn \setminus cn) / (s / n), n, (n \setminus s) / n \vdash n} [\text{略}] \quad \frac{n, (n \setminus s) / n, n \vdash s}{n, (n \setminus s) / n \vdash s / n} [\text{略}]}{n / cn, cn, (cn \setminus cn) / (s / n), n, (n \setminus s) / n \vdash n} [/L]$$

the student who Pat saw

ここで、右上のシーケントは *Pat saw Kim* のような完全な他動詞文だが、目的語は仮定として置かれたもので、 $[/R]$  規則により右辺に移動される。その結果生じた型  $s / n$  は目的語の欠けた節を表す。Who の型は  $s / n$  を引数にとる高階関数であり、これによって正しく関係節が派生される。<sup>8)</sup>

型論理文法のもう一つの利点は、合成的意味論との対応である。技術的詳細は省くが、証明と  $\lambda$  項の Curry-Howard 対応により、上記の証明から直ちに次の  $\lambda$  項が得られる：

$$x_{\text{the}}(x_{\text{who}}(\lambda y. x_{\text{saw}}(y)(x_{\text{Pat}}))(x_{\text{student}}))$$

各自由変数を、語彙の具体的な意味表示に置き換えれば、例えば次のような意味表示が得られる：

$$\lambda x. \text{student}(x) \wedge \text{saw}(x)(\text{Pat})$$

この意味表示に型が付与できることは、言語表現の統語上の型とその意味の型の対応（準同型）を想定することで保証される。単純な外延意味論を想定した場合この準同型  $h$  は以下のようなになる：

$$h(s) = t; \quad h(n) = e; \quad h(cn) = e \multimap t;$$

$$h(b \setminus a) = h(a / b) = h(b) \multimap h(a) \quad \text{for any } a, b$$

たとえば *saw* の統語型  $(n \setminus s) / n$  は意味型  $e \multimap (e \multimap t)$  に対応する。

### 3 これからの統語理論として

本節では、型論理文法が、従来の他の統語理論と統語構造についての基本的な直観を共有していることを見たとえ、主流生成文法に向けられている批判点や LLM の登場がもたらした課題に、どのように答えているのか、あるいは答えるのかをみる。

8) 実際には目的語関係節において空所の位置は必ずしも節の右端ではない（例：*who Pat saw \_\_ yesterday*）ため、この分析では不十分である。似た問題は量子子などの分析にも生じ、総称して不連続性 (discontinuity) の問題とも言われる。この点に対処した  $L$  の拡張は多数ある ([19, 20, 21] など)。



響の限界を示しているというのが、型論理文法からの見立てになる。この観点は、LLM や人間における統語と意味の統合を再検討する起点になりうる。また、分布意味論と型論理に基づく合成的意味論との統合も追求されている [28].

### 3.5 逐次処理との親和性

人間の言語産出・理解は常に時間の流れの中で逐次的に起こり、統語処理も例外ではない [29]. ゆえに、統語理論は逐次的な言語使用の理論に埋め込み可能であるべきで、それは元々 Chomsky も認識していた [30]. しかし、歴史的な経緯もあり、主流生成文法は逐次処理への埋め込みを事実上放棄した。そのことは言語学者からしばしば批判されてきた [31, 7]. 統語木を「下から」組み上げる形で定義された主流生成文法を「左から右」に処理する難しさは、近年の取り組みが示している [32, 33].

型論理文法にはそうした困難はない。逐次的処理の途中の断片に対しても型を与えられるからである。例えば、関係節の分析でみたように

$$n, (n \setminus s) / n \vdash s / n$$

は証明可能なシーケントである。これは、例えば *Pat saw* まで入力され、目的語を待っている状態を表しているとみることができる。統語計算ができるので意味表示も与えられる。別の言い方をすれば、シーケントは状態遷移規則とみなせる。

研究史上、逐次処理との親和性を強調してきたのは CCG であり、実証研究でも CCG が他の文法とは独立に人間の大規模な行動・生理データを捉えられることが示されてきた [34, 35, 36]. CCG は、Lambek 流の型論理文法で証明可能な定理のサブセットを規則として定めたものとみることができる [37]. 興味深いことに、サブセットの選択により、直ちには妥当な型が付与できないような文の断片が生じる。例えば次の断片はいずれも、CCG の標準的な規則では妥当な型を付与できない：

the student who Pat  
花子が太郎の

このような CCG の逐次性の限界は批判的に見られることもあるが [38], 磯野ら [35, 36] は、型の付与を情報圧縮とみなし、型が付与できない断片は圧縮できないため作業記憶に負荷をもたらすと仮定すると、英語や日本語の自然な文章における人間の単語あたりの読み時間の変動を予測できることを実証的

に示している。つまり、CCG が仮定する型付与に対する限界が、人間のリアルタイムの言語処理能力の限界を捉えているとみられる。裏を返せば、L のような強力な型論理文法は、それが許容するあらゆる推論が人間に可能ではないかもしれないが、人間の限界を探る上でも重要となる、見通しのよい体系を与えているといえる。なお、人間の言語処理の確率的な側面との統合も今後の重要な課題である。

### 3.6 進化・獲得の問題との接続可能性

主流生成文法の言語能力の生得性についての主張は、LLM の登場で改めて批判的になっている。統語を型論理とみなすと、統語の進化や獲得について、他の認知能力の進化や獲得について形式論理の観点を交えて検討してきた研究とも接続できる可能性がある。例えば、Kemp ら [39] は、人間が持つ直観的な生物学や物理学を形式的に記述しようとするれば述語論理が必要となることを指摘した上で、一階述語論理がデータから階層ベイズモデルによって獲得できる可能性を示唆している。

進化・獲得を巡っては、用法基盤モデルが主流生成文法と大きく異なる立場を提示してきた [40, 41]. 同モデルが強調する、学習過程や個別言語にみられる特異的な構文を、型論理文法も扱えることを簡単に示しておく。By and large のような固定した表現は、単に個別文法の語彙に登録しておけばよい。Give NP the cold shoulder のような非連続の固定した表現は、the cold shoulder に専用の型  $n_{ics}$  を与え give の項に指定すれば実現できる。また、固定した表現のない構文、たとえば二重目的語構文 (*[Pat gave] Kim the book*) に固有の意味は型論理文法ではあまり議論されないが、構文として分析したければ、発音されない要素を構文の主要部として置くことで分析可能である ( $sc$  は小節 (small clause) の型) :<sup>9)</sup>

$$\underbrace{(n \setminus s) / sc}_{gave}, \underbrace{n}_{Kim}, \underbrace{(n \setminus sc) / n}_{\emptyset}, \underbrace{n}_{the\ book} \vdash n \setminus s$$

## 4 おわりに

統語を一種の型論理とみなす枠組みが、LLM 以降の時代の統語理論として有効であると論じた。型論理文法に目を向ける人が増えることを願う。

9) このような分析は主流生成文法でみられる [42]. 型論理文法では、空文字の型は  $\bullet$  の単位元として一意に定まってしまうため、このような分析は通常行われないが、型の解釈先の抽象的な単語列と、実際に発音される音韻表示を別の表示レベルとして扱えば論理体系を維持したまま実現可能である。

## 謝辞

LLMと言語学をめぐる論争については、占部雄大氏・鴨田豪氏・窪田悠介氏・中石海氏・山本悠士氏・横井祥氏（五十音順）との勉強会での議論から多くの示唆を得た。また、本稿の原型となった Evidence-based Linguistics Workshop 2025（日本）および Computational Psycholinguistics Meeting 2025（オランダ）での発表でも、参加者との議論から示唆を得た。本研究は JSPS 科研費 JP25K22996 の助成を受けている。

## 参考文献

- [1] Steven Piantadosi. Modern language models refute Chomsky’s approach to language. In Edward Gibson and Moshe Poliak, editors, **From fieldwork to linguistic theory: A tribute to Dan Everett**, pp. 353–414. Language Science Press, 2024.
- [2] Richard Futrell and Kyle Mahowald. How linguistics learned to stop worrying and love the language models. **arXiv preprint arXiv:2501.17047**, 2025.
- [3] Daniel C. Dennett. Real patterns. **Journal of Philosophy**, Vol. 88, No. 1, pp. 27–51, 1991.
- [4] Noam Chomsky. **The Minimalist Program**. MIT Press, 1995.
- [5] Noam Chomsky, T. Daniel Seely, Robert C. Berwick, Sandiway Fong, M. A. C. Huybregts, Hisatsugu Kitahara, Andrew McInerney, and Yushi Sugimoto. **Merge and the Strong Minimalist Thesis**. Cambridge University Press, 2023.
- [6] Robert D. Borsley and Stefan Müller. HPSG and Minimalism. In Stefan Müller, Anne Abeillé, Robert D. Borsley, and Jean-Pierre Koenig, editors, **Head-Driven Phrase Structure Grammar: The handbook**, pp. 1253–1329. Language Science Press, 2021.
- [7] Ivan A. Sag and Thomas Wasow. Performance-compatible competence grammar. In Robert D. Borsley and Kersti Börjars, editors, **Non-transformational syntax: Formal and explicit models of grammar**, pp. 359–377. Blackwell, 2011.
- [8] Cristiano Chesì. Is it the end of (generative) linguistics as we know it? **Italian Journal of Linguistics**, Vol. 37, No. 1, pp. 131–144, 2025.
- [9] 戸次大介. 合理主義言語学における部分構造論理. 人工知能学会誌, Vol. 27, No. 3, pp. 304–309.
- [10] Kazimierz Ajdukiewicz. Syntactic connexion [translated by h. weber]. In Storrs McCall, editor, **Polish logic: 1920–1939**, pp. 207–231. Oxford University Press, 1935/1967.
- [11] Joachim Lambek. The mathematics of sentence structure. **The American Mathematical Monthly**, Vol. 65, No. 3, pp. 154–170, 1958.
- [12] Joachim Lambek. Structure of language and its mathematical aspects. In **Proceedings of Symposia in Applied Mathematics 12**, 1961.
- [13] Mark Steedman. **Surface structure and interpretation**. MIT Press, 1996.
- [14] Daisuke Bekki. Combinatory Categorical Grammar as a substructural logic JSAI-isAI 2010. In **New Frontiers in Artificial Intelligence**, 2011.
- [15] 戸次大介. 日本語文法の形式理論. ころしお出版, 2010.
- [16] 谷中瞳. ことばの意味を計算するしくみ 計算言語学と自然言語処理の基礎. 講談社, 2024.
- [17] Michael Moortgat. Categorical Type Logics. In Johan van Benthem and Alice ter Meulen, editors, **Handbook of logic and language**, pp. 93–177. MIT Press, 1997.
- [18] Mark Steedman. Dependency and coordination in the grammar of Dutch and English. **Language**, Vol. 61, No. 3, pp. 523–568, 1985.
- [19] Glyn V. Morrill, Neil Leslie, Mark Hepple, and Guy Barry. Categorical deductions and structural operations. In Guy Barry and Glyn V. Morrill, editors, **Studies in Categorical Grammar (Edinburgh Working Papers in Linguistics)**. 1990.
- [20] Glyn V. Morrill. **Type logical grammar**. Kluwer, 1994.
- [21] Yusuke Kubota and Robert D. Levine. **Type-logical syntax**. MIT Press, 2020.
- [22] Mark Steedman. On internal merge. **Linguistic Inquiry**, Vol. 57, No. 1, pp. 113–171, 2026.
- [23] Dirk Roorda. **Resource logics: Proof-theoretical investigations**. PhD thesis, University of Amsterdam, 1991.
- [24] Glen V. Morrill. **Categorical Grammar: Logical Syntax, Semantics, and Processing**. Oxford University Press, 2011.
- [25] Chris Collins and Edward P. Stabler. A formalization of Minimalist syntax. **Syntax**, Vol. 19, No. 1, pp. 43–78, 2016.
- [26] Noam Chomsky. **Syntactic structures**. Mouton, 1957.
- [27] Kristina Gulordava, Piotr Bojanowski, Edouard Grave, Tal Linzen, and Marco Baroni. Colorless green recurrent networks dream hierarchically. In **NAACL:HLT, Vol. 1 (Long Papers)**, 2018.
- [28] Dmitri Kartsaklis. Compositional operators in distributional semantics. **Springer Science Reviews**, Vol. 2, pp. 161–177, 2014.
- [29] M. K. Tanenhaus, M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy. Integration of visual and linguistic information in spoken language comprehension. **Science**, Vol. 268, No. 5217, pp. 1632–1634, 1995.
- [30] Noam Chomsky. **Aspects of the theory of syntax**. MIT Press, 1965.
- [31] Joan Bresnan and Ronald Kaplan. Introduction: Grammars as mental representations of language. In Joan Bresnan, editor, **The mental representation of grammatical relations**, pp. i–lii. MIT Press, 1982.
- [32] Tim Hunter. Left-corner parsing of Minimalist Grammars. In Robert C. Berwick and Edward P. Stabler, editors, **Minimalist Parsing**. Oxford University Press, 2019.
- [33] Miloš Stanojević and Edward P. Stabler. A sound and complete left-corner parsing for Minimalist Grammars. In **CogACLL2018**, 2018.
- [34] Miloš Stanojević, Jonathan R. Brennan, Donald Dunagan, Mark Steedman, and John T. Hale. Modeling structure-building in the brain with CCG parsing and large language models. **Cognitive Science**, Vol. 47, No. 7, p. e13312, 2023.
- [35] Shinnosuke Isono. Category Locality Theory: A unified account of locality effects in sentence comprehension. **Cognition**, Vol. 247, p. 105766, 2024.
- [36] Shinnosuke Isono, Kohei Kajikawa, Yohei Oseki, and Masayuki Asahara. Modeling memory effects in a head-final language with category locality. **PsyArXiv preprint PsyArXiv:aj6my\_v1**, 2025.
- [37] Jason Baldridge. **Lexically specified derivational control in Combinatory Categorical Grammar**. PhD thesis, University of Edinburgh, 2002.
- [38] Vera Demberg. Incremental derivations in CCG. In **TAG+11**, pp. 198–206, 2012.
- [39] Charles Kemp, Noah D. Goodman, and Thomas L. Griffiths. Bayesian inference over logical representations. In Thomas L. Griffiths, Nick Chater, and Joshua B. Tenenbaum, editors, **Bayesian Models of Cognition**, pp. 423–435. MIT Press, 2024.
- [40] Adele E. Goldberg. **Constructions: A Construction Grammar Approach to Argument Structure**. University of Chicago Press, 1995.
- [41] Michael Tomasello. **Constructing a language: a usage-based theory of language acquisition**. Harvard University Press, 2003.
- [42] Liina Pyllkkänen. **Introducing arguments**. MIT Press, 2008.