

新進研究者 Research Notes

日本科学哲学会

科学基礎論学会

第2号（2019年）

比較言語学における局所的最節約法 —分岐学との対比を通じて—	吉田哲郎	1
私的言語の可能性と不可能性	小森勇輝	9
世界3の自律性は言語的実体と衝突するのか	池田健人	18
予測符号化の哲学的含意と過学習の問題	柴田翔平	27
根元的解釈の基礎について	中谷内悠	36
マイノング主義論理におけるボトムアップの観点	小関健太郎	44

JAPANESE STUDENT RESEARCH NOTES OF PHILOSOPHY OF SCIENCE

Philosophy of Science Society, Japan

Japan Association for Philosophy of Science

No. 2

Contents

Tetsuro Yoshida: Local Parsimony in Comparative Linguistics: as Contrasted with Cladistics	1
Yuki Komori: The Possibility and Impossibility of Private Language	9
Kento Ikeda: Does the Autonomy of World 3 Conflict with Linguistic Entities?	18
Shohei Shibata: Philosophical Implication of Predictive Coding and the Problem of Overfitting	27
Yu Nakayachi: On The Basis of Radical Interpretation	36
Kentaro Ozeki: A Bottom-up View on Meinongian Logics	44

比較言語学における局所的最節約法
—分岐学との対比を通じて—

Local Parsimony in Comparative Linguistics:
as Contrasted with Cladistics

吉田 哲郎

Abstract

Some researchers in comparative linguistics reconstruct a three-vowel system for early Proto-Indo-European. This vowel system is doubtful from the standpoint of language universals. The aim of this paper is to clarify the background of this conflict. I point out that the cause of the problem is the method “parsimony about the number of vowel phonemes”. Such parsimony, permitted in the name of “Ockham’s razor”, works well for reconstructing the earliest stage of proto language. As contrasted with cladistics, parsimony in comparative linguistics is used only in one language family. This locality is based on a cautious attitude toward phylogenetic relationships among language families.

(1) 研究テーマ

比較言語学で用いられる最節約法の独自性と、その最節約法が許容される背景。

(2) 研究の背景・先行研究

比較言語学(comparative linguistics)は、互いに類似する諸言語を観察し共通祖先の仮説に基づいて、共通祖先である言語（祖語）の姿を明らかにする歴史的な研究である。こうした祖語の復元の試みは、再建と呼ばれる。もはや観察できない過去の言語である祖語を再建するため、いくつかの手法が存在する。そのうち、比較方法と呼ばれる手法は、生物学における分岐学(cladistics)の方法論と基本的に同一とみなされている（三中 1997 p.87）。

比較言語学における祖語の再建は、特にインド＝ヨーロッパ語族（印欧語族）に対して成功を収めてきた。そして、印欧語族の祖語すなわち印欧祖語がどのような母音体系を持っていたかについて、音韻論の領域において仮説が提示されてきている。たとえば、古典的学説は、印欧祖語が、i, e, a, o, u の5母音を有していたとする。

研究者の中には、印欧祖語を時間軸に沿って動的に変化するものと捉えた

うえで、その最古の状態まで再建ができると考える者もいる。たとえば、神山(2003)は、喉音理論(laryngeal theory)と呼ばれる理論を採用したうえで、印欧祖語の最古の状態では母音は e ひとつだけであったと述べている。神山は、i,u を母音に含めていないが、i,u を母音に含めるならば¹、神山の説は i,e,u を再建する印欧祖語最古期 3 母音体系説ということになる。

ところが、この 3 母音体系は、言語普遍性(language universals)の研究からは支持されない。言語普遍性の研究とは、特定の語族に偏ることなく世界の諸言語を広く均等に観察してデータを蓄積し、言語の普遍性もしくは傾向を把握しようとする言語類型論的な研究である²。たとえば、世界の諸言語がどのような音素を有しているかについてのデータベースとして、UPSID (UCLA Phonological Segment Inventory Database)があり、20 の語族から 317 の言語が選ばれて収録されている。これによれば、5 母音体系を有する言語が 109 あり、317 の言語に占める割合は約 34%で最頻である。次いで 6 母音体系(約 19%)、7 母音体系(約 14%)と続く。3 母音体系は 19 言語で約 6%である。3 母音体系の中で最も多いのは i,a,u の体系である(18 言語)。i,e,u の 3 母音体系は、UPSID に存在しない³ (Schwartz et al. 1997)。実例がないのであるから、喉音理論が提示する i,e,u という 3 母音体系が現実に人間言語として成立しうる母音体系なのか不明である。データベースにより詳細は異なるにせよ、印欧祖語最古期 3 母音体系説は、言語普遍性の観点から再考の余地を指摘されている (Hock1991,p.572)。

実際に、こうした再建を認めておくことはできないとする論者がいる。たとえば、コムリーは、歴史的な言語研究が言語普遍性の研究と整合すべきと考えている。彼は、ある言語タイプが実際に起こりえないということが言語普遍性の研究によって示唆された場合、そのタイプの言語を祖語として設定するような再建は破棄されなくてはならないと述べる (Comrie1981)。

他方で、そこまでの強い主張をしない論者もいる。たとえばホックは次のように述べる。言語類型論がすべての言語を視野に入れることは困難である。記録されることなく滅びた言語もある。言語類型論がそうした言語をカバーできない以上、比較言語学における成果が言語類型論の知見と食い違っているというだけで比較言語学的な成果を捨て去るのは、早計である (Hock1991,p.626)。

結局のところ、この状況は問題であると認識されつつも、そのまま解決されることなく今に至るようである。事実、比較言語学も、言語普遍性の研究(言語類型論)も、ともに科学的な手法のもとこれまで成果を積み上げてきており、単純にどちらかが間違いであると断じることはできない。問題の解

決に向けた準備としてまず必要なのは、方法論が異なることを理解し、それぞれの方法論を成立させている背景事情を理解することだろう。

本稿では、比較言語学の方法論に焦点を当て、研究手法の事実関係の整理を主に行う。そして、この母音体系の問題の原因は、比較言語学で用いられる「母音音素数に関する最節約法」にあると主張する。この最節約法は、生物学における分岐学が用いる最節約法とは異なるタイプのものである。本稿の目的は、分岐学を参照することで比較言語学に用いられる最節約法の独自性を明らかにし、その最節約法を許容する背景に系統関係の認識に関する慎重な姿勢があることを指摘することにある。

まず、分岐学と比較言語学で共通する最節約法の考え方を、祖先形質の推定の思考例を通じて概観しよう。次の例1は、ある系統樹が与えられたときの祖先形質状態を末端で観察される形質から最節約的に推定する問題である (Sober 2015, p.163 に基づく)。以下の例では【前提条件】、【問い】及び【回答】の3項目をたてることにする。

【例1：前提条件】子孫種 $D_1 \sim D_{10}$ は、すべて共通祖先 A から分化して生じたとする。今、これらの子孫種のある形質に着目している。10種の子孫種のうち、 D_2 のみが、その形質に関し状態1を示している。残りの9種はすべて、その形質に関し状態0を示している (図1)。

【例1：問い】 A における当該形質の状態は何であったと推定すべきか？

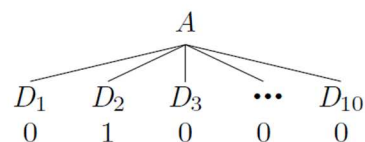


図1

【例1：回答】最節約法により、 A の形質状態は0であったと推定できる。 D_2 に至る枝で形質変化 $0 \rightarrow 1$ を1回措定するほうが、形質変化 $1 \rightarrow 0$ を他の9本の枝でそれぞれ措定するよりも、最節約的である。

この回答は、形質変化の回数が9回となるように考えるよりも、1回となるように考えるほうが望ましいという、「変化回数に関する最節約法」に基づいている。比較言語学者も同様の思考原理を採用して祖語の母音に関する推定を行う。次に、例1を、比較言語学における状況に置き換えてみると、次の例2になる。

【例2：前提条件】子孫言語 $L_1 \sim L_{10}$ は、すべて祖語 L から分化して生じた

とする。今、これらの子孫言語の母音体系に着目している。10の子孫言語のうち、L₂のみが、母音 V を持たない。残りの9つの言語はすべて、母音 V を有している。

【例2：問い】祖語 L における母音は、どのようであったと推定すべきか？

【例2：回答】最節約法により、L は V を有していたと推定できる。V が消失する変化を1回措定するほうが、V が生じる変化を9回措定するよりも、最節約的である。

この例2は、例1と同様、「変化回数に関する最節約法」を用いている。この再建手法は、比較言語学において「比較方法（比較法）」と呼ばれている。また、この節約原理は、オッカムの剃刀と表現される（Hock 1991, p.538）。このように、「変化回数に関する最節約法」は、分岐学と比較言語学で共通する手法である。

(3) 筆者の主張

上述の例1と例2における「変化回数に関する最節約法」は、分岐学においても比較言語学においても用いられるものである。次に、比較言語学で独自に用いられる最節約法の例を挙げることにする。以下の例3は、その最節約法の例であり、喉音理論をモデルとしている。

【例3：前提条件】今、言語 L の母音体系に着目している。L は、e, a, o という3母音体系を示している。その3母音のうち、e の出現頻度が最も高い。

【例3：問い】L は、過去にいくつの母音を持つ体系だったと推定できるか？

【例3：回答】最節約法により、L の過去の母音体系は e のみの1母音であったと推定できる。e のみを措定するほうが、e, a, o を措定するよりも、最節約的である。

この最節約法の要は、母音の数に着目して、過去の母音体系を推定する点にある。その際に、観察される個数が最も多い母音音素を祖先にするという発想に基づいている⁴。つまりこの最節約法で節約的にする対象は、変化の回数ではなく、母音の数なのである。よって、この最節約法は「母音音素数に関する最節約法」と呼べるだろう。これは先の例で見た「変化回数に関する最節約法」とは異なるが、現に比較言語学者が用いる最節約法のひとつである。例えば、ホックは、再建で音素の数について最節約的にすべきとの考えを示している。そして、この思考方法もまたオッカムの剃刀と表現される

(Hock 1991, p.538)。つまり、比較言語学では、オッカムの剃刀の名のもとに、2つのタイプの最節約法が認められているのである。

この最節約法は、言語普遍性の観点から不自然な母音体系を作り出してしまうというのが、本稿で取り上げている問題である。この最節約法が適用された結果、過去の母音体系は e だけの 1 母音体系となる。i, u を母音に含めるならば、i, e, u の 3 母音体系ということになる。そして、言語類型論的なデータベースにこの体系が見当たらないことは、既に述べたとおりである。果たして、言語として現実に有り得る母音体系なのか疑わしい。

また、例 1・2 の最節約法と、例 3 の最節約法は、相反する関係にある。過去の母音の数を減らすことで、変化のプロセスが増える。たとえば、1 母音が 3 母音になったとする場合、母音音素の分岐変化を措定することが必要となり、変化回数に関する最節約法の徹底を妨げてしまうのである。

さらに例 3 の最節約法は、生物に適用してみても、奇妙な推論を招いてしまう。たとえば、「ヒグマとシロクマでは、ヒグマの個体数が多く観察されるので、ヒグマがシロクマの祖先である」という主張を考えてみよう。この主張は、現在多く観察される方が祖先であることを前提にしている。しかし、祖先の方が多く残存するのだろうか。そうとは限らないであろう。

ソーバーは、種 A と種 B が類似しており、A の化石が B の化石より多く見つかるからといって、A は B の祖先であると主張できるわけではないとし、そう主張するための条件として、①化石証拠がほぼ完全であること、②種内進化が十分に漸進的であること、③A と B の共通祖先である種 C が存在しないこと、以上 3 つを挙げている (Sober 1988, p.20)。

比較言語学において、これらの条件が満たされているとは考えられない。まず①について、私たちが観察できる化石がわずかな数である (化石記録の不完全性) のと同様、私たちが観察できる過去の言語資料もわずかな数である。過去の言語を文字で記した紙や石碑、粘土板が我々の時代まで残る保証はなく、風化したり焼失したり破壊されたりしたのであろう。こうした不完全な言語資料の中で e が頻出するのは、たまたま残存した資料が偶然に持つ特徴にすぎないかもしれない。こうした中で引き出される結論の妥当性は、疑わしいであろう。そして、②について、母音の変化の速さを一定だとする仮定を置くのは困難である。③について、共通の源となる他の母音が存在した可能性を排除する方法はないように思われる。

以上の問題があるにも関わらず、比較言語学では「母音音素数に関する最節約法」が使用されている。その理由は、分岐学における外群(outgroup)に相当する概念が比較言語学にないため、祖語の再建及び祖語のより古い状態

の再建を、内群(ingroup)から行わなければならないからである。

分岐学の外群比較法という手法によると、いま調べようとしている群（内群）に対して近縁であると仮定される群（外群）に基づいて、内群の形質方向性が判定される（三中 1997,p.112）。ここで重要なのは、定義上、外群は内群に系統的關係があるという点である。例えば以下の図 2（Sober2015,p.158 を改変）において、 O_1, O_2, \dots, O_n は外群であり、

A を祖先とする内群と系統關係にあると仮定されている。これら外群を、例 1 における $D_1 \sim D_{10}$ のように扱って最節約法を適用することで、内群の祖先（図 2 の A）の形質状態を、いわば「外側から」決めることが可能である。

その次に、内群の分岐図が推定される⁵。

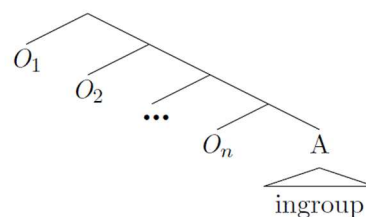


図 2

こうした外群比較法を正当化する根拠もまた、分岐学的最節約法、すなわち変化回数に関する最節約法にある（Sober2015, p.158）。したがって、分岐学における最節約法は、変化回数に関するものだけで一貫しており、内群だけでなく外群も含めて大域的(global)に変化回数を最節約的にするという思考原理なのである。これが可能なのは、いま研究対象としている内群を含み、すべての生物をカバーする大きな系統樹が仮定されているためである。つまり生物学では、すべての現生生物が DNA を遺伝物質として利用している等の普遍的相同性の存在によって、すべての生物が同じ共通祖先に遡ると考えられている。

これに対して言語学では、すべての言語が同じ共通祖先言語に遡るとは仮定されていない。ある語族の外の言語は、当該語族と系統的關係がない言語とされている⁶。例えば、印欧語族とシナ＝チベット語族が系統關係にあるとは考えられていない。これは、比較言語学ではある語族の外群に相当する概念自体が存在しないということである。比較言語学が、現存する過去の言語資料が乏しい中で、ある語族と他の語族との系統關係を安易に措定しないのは、認識論的には慎重な姿勢であるだろう。

外群に相当する概念がないので、研究対象とする語族の外側から祖語の特徴を推定することができず、外群を含め大域的に変化回数を最節約的にするという発想も生じ得ない。比較言語学では、ひとつの語族の中から、いわば「内側から」、当該語族の祖語を再建するほかないのである。そして、再建された祖語の最古の状態の母音体系を知りたいと思うならば、その祖語自体から最古の状態を推定する手法に頼るしかない。その手法とは、「母音音素数に関する最節約法」である。例 3 で見たように、この最節約法はひとつの言語

だけから、さらにその過去を推定できるので、祖語だけから、さらにその祖語の最古の状態を再建できる。そして、この最節約法もまた、「オッカムの剃刀」という名のもとに正当化され使用されてきた。

結論を分岐学と対比的に述べよう。分岐学の最節約法は、外群を含み大域的に変化回数を最節約的にする。これに対して比較言語学の最節約法は、他の語族を参照しないという意味において局所的(local)である。その局所性ゆえに、この最節約法は当該語族内だけで完結して使用できるが、局所解を提案してしまう。このことが、本稿冒頭に示した、言語普遍性の研究との不整合を招いているのである。

(4) 今後の展望

以上のように比較言語学の最節約法の背景を理解すれば、母音体系の問題の解決の糸口も見えてくる。たとえば、もし比較言語学でも外群概念の仮定が認められれば、大域的に変化回数を最節約的にする方針のもと、他の語族のデータも考察に含めることができる。そして、母音音素数に関する最節約法を使わないのであれば、不自然な母音体系が祖語の最古期に措定されることもなくなるだろう。

ただし、外群の概念を得るためには、語族と他の語族との間に系統関係があることを仮定しなくてはならない。これは比較言語学が慎重にも採用してこなかった仮定である。しかし、資料の不完全により語族間の系統関係が不明だからといって「系統関係がない」とまで主張しなくてはならないわけではない。具体的な系統関係は分からないとしても、何らかの系統関係がある可能性だけでも仮定できないだろうか。今後はこうした、系統関係の仮定に関する認識論的な議論を深めなくてはならないだろう。

注

1. 神山(2003)では i 及び u は母音とみなされていないので、印欧祖語の最古の状態では母音は e だけだったと記述されている。しかし言語学の多くの分野で i 及び u は母音とみなされている。本稿では後述する言語普遍性の研究と比較可能にするために、i 及び u を母音とみなすことにする。

2. UPSID には 1 母音体系と 2 母音体系も存在せず、母音数が最少の体系は 3 母音体系である (Schwartz et al.1997)。

3. 本稿では言語類型論と言語普遍性の研究をほぼ同義で用いる。

4. 喉音理論では、e, a, o のうち、最頻出の母音 e が祖先的な母音と見なされ、a と o の多くを、e と喉音と呼ばれる音との組み合わせにより後代に生

じた子孫的な母音と見なす。この考え方は、ソシュールに遡る (Saussure 1878, p.135)。

5. ただし、外群に最も多い形質状態が祖先形質状態だと判断されるわけではない。内群に最も近い祖先形質状態が、内群の祖先形質状態を決める。本稿ではこの詳細は捨象する。なぜなら本稿で問題とするのは外群の有無であり、外群の存在を前提としたうえでの議論は、本稿に影響しないためである。

6. たしかにこれまで印欧祖語を超えてより高いレベルで「Nostratic 祖語」や「世界祖語」といった祖語を再建することで、語族間の系統関係を立証しようとする動きがあったが、過去の言語資料の不完全性による我々の限られた知識から、そうした再建は成功していない (Clackson 2007, p.20)。

(5) 参考文献

- 神山孝夫, 2003, 「印欧祖語のアップラウトと文法構造の発達」, 『大阪外国語大学論集』 27.
- 三中信宏, 1995, 「分岐分類学に基づく形質進化の最節約復元」, 『種生物学研究』 19, 39-49.
- 三中信宏, 1997, 『生物系統学』, 東京大学出版会
- Clackson, J., 2007, *Indo-European Linguistics: An Introduction*, Cambridge University Press.
- Comrie, B, 1981, *Language Universals and Linguistic Typology: Syntax and Morphology*, Basil Blackwell. (邦訳: バーナード・コムリー 『言語普遍性と言語類型論 — 統語論と形態論 —』 松本克己・山本秀樹訳 ひつじ書房 1992年)
- Hock, H. H., 1991, *Principles of Historical Linguistics*, Mouton de Gruyter.
- Saussure, F., 1878, *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*, Teubner.
- Schwartz, J.L., Boë, L.J., Vallée, N., and Abry, C., 1997, “Major trends in vowel system inventories”, *Journal of Phonetics*, 25, 233-253.
- Sober, E., 1988, *Reconstructing the past — Parsimony, Evolution, and Inference*. MIT Press. (邦訳: エリオット・ソーバー 『過去を復元する』 三中信宏訳 勁草書房 2010年)
- Sober, E., 2015, *Ockham's razors: a user's manual*, Cambridge University Press.

(放送大学)

私的言語の可能性と不可能性

The Possibility and Impossibility of Private Language

小森 勇輝

Abstract

In this paper, I will focus on Wittgenstein's private language argument, especially on the case of keeping a diary about the recurrence of sensations. Many philosophers have claimed that the sign "E", which refers to speaker's private sensation, couldn't be understood by the others, for it lacks criteria of correctness. To this, I will reply that "E" *can* be understood as a sign of 'sensation', which belongs to our common language. With this, the possibility to use "E" without criteria arises. However, since the identification of 'sensation' doesn't succeed, "E" *cannot* be replaced by common language. I shall argue that this contradiction is essential to private language argument.

(1) 研究テーマ

本研究テーマは、後期ウィトゲンシュタイン哲学における「私的言語論」である。とりわけ、「感覚日記」の事例を参照しながら議論を展開していく。

(2) 研究の背景・先行研究

A・J・エイヤーと R・リースの論争を皮切りに (cf. Ayer 1954, Rhees 1954)、私的言語 (private Sprache) の可能性と不可能性をめぐる L・ウィトゲンシュタインの議論は、様々な研究者たちの関心を引きつけるようになった。とりわけ S・A・クリプキが、一般に「規則遵守についての考察 (rule following considerations)」と呼ばれている一連の断章から「共同体説」を導き出して以来、『哲学探究』(以下『探究』または PU と略記)における「規則には『私的に』従うことはできない」(PU・202) という所見は、私的言語の不可能性を示すための論拠として注目を集めてきた (cf. Kripke 1982)。いわゆる「私的言語論 (private language argument)」は、第 243 節以降に展開されると言われているが、クリプキの考えでは、「真の『私的言語論』は第 243 節に先立つ諸節において見出されるべきものである。事実、第 202 節において、『私的言語論』の結論はすでに明示的に述べられている」(ibid. p.3 強調原文)。

とはいえ、規則遵守についての一連の考察から「『私的な』規則遵守は不可

能である」という一般的な帰結が導かれるとしても、果たしてそれによって私的言語の不可能性までもが示されることになるのだろうか。実際、「私的言語論」の中心を占めている「私的 (private)」という概念には、「共同体から切り離されて考えられた個人」(ibid. p.110 強調原文) という逸脱的なニュアンスよりも、「それを感じている「私」にとって不可謬であるような感覚の私秘性ⁱ⁾」という、一人称権威を象徴する意味合いの方が、一段と強く含まれているように思われる。このような見立てのもと、本稿は私的言語についての一連の議論を、「規則遵守論」の「系」とは看做さずに、あくまでも感覚の言語ゲームにまつわる独立の主題として考察していく。

さて、私的言語とは、その言語の話し手の「直接的で私的な感覚を指示する」(PU-243) ためだけに用いられる言語である。それゆえ、「他人はこの言語を理解することはできない」(ibid.)。本稿は、第 243 節におけるこの「定義」を念頭に置きながら、第 258 節以降に登場する「感覚日記」の事例に焦点を当てていく。これは繰り返し生起する「感覚 (Empfindung)」を「名」に結びつける行為として記号「E」を日記に書き込む、という事例であり、長らく「私的言語論」の中心を占めると言われてきた。この事例の最大の特徴は、日記に書き込まれる「E」の「定義」が述べられない点にある。

次のような場合を想像しよう。私はある感覚が繰り返し起こることについて、日記をつけようと考えた。そこでその感覚を記号「E」に結びつけることにして、その感覚を持った日には必ずカレンダーに記号「E」を書き込む。——私がまず言っておきたいのは、その記号の定義を述べることはできないということだ。——だが、私は自分自身には一種の直示的定義 (hinweisende Definition) を与えることができる！——どうやって？ 私は感覚を指し示すことができるのか？ ——普通の意味ではできない。だがその記号を口にしたり書いたりして、自分の注意をその感覚に向ける——いわば心の中でその感覚を指し示すのだ。——しかし、何のためにそんな儀式 (Zeremonie) を？ というのも、それはそういった儀式にしか見えないからだ！ 定義は、記号の意味を確定するのに役立つ。——ところで、確定することは、まさに注意を集中することによって行われる。それによって、記号と感覚の結びつきを自分に刻みつけるのだから。(PU-258)

記号「E」に結びつけられる感覚とは、自然な表出 (叫び声、顰めっ面、額の脂汗…) からは切り離されて起こる感覚のことである (cf. PU-256, 257)。例えば「痛み」のような通常感覚語であれば、自然な表出は「痛み」の使用の正しさや有意味性を保証する「規準 (Kriterium, criterion) ⁱⁱ⁾」として機能することになる。これに対して「E」が日記に書き込まれる場合には、この

記号の書き込みを正当化する規準は見受けられない（「E」の「定義」を述べるができない所以である）。したがって第 258 節は、公共的に観察可能な規準が存在しない状況で書き込まれる「E」が、そもそも「言語」として成立するのかどうかを問題にしているのである。

これまで多くの論者たちは、私的言語の不可能性を、「E」の適用に規準が伴わない点に見出してきた（cf. Malcolm 1954 p.97, Hacker 1990 pp.118-123, Bertolet 1995 p.647, Glock 1996 p.311）。この手の批判は次のように要約される——「痛み」のような通常的感覺語は、自然な表出を介して感覺と結びついたときに、初めて意味を獲得する。そして、このような規準の設定によって、その後の感覺語の使用の正・誤が判定される。すなわち最初に「痛み」の定義が設けられ、その定義に照らしてその都度の「痛み」の使用に「正しい（richtig）」とか「正しくない（unrichtig）」といった評価が下されるのである。ところが、「感覺日記」の事例においては、このような手続きは踏まれない。「E」の適用を正当化するには、「この感覺を「E」に結びつけたはずである」というかつての記憶や、自分の心の中で繰り広げられる想像に頼るしか方法がないため、「E」の適用が本人にとって正しいと思われるだけで、その適用は自動的に正しいことになってしまう。かくして、「感覺日記」の事例においては、公共的な規準が欠けているために、「正しい（richtig）」と「正しく思われる（richtig erscheinen）」との区別が設けられなくなる。

——「私はそれを自分に刻みつける」とは、その過程を経ることによって将来その結びつきを正しく思い出すことができる、ということに過ぎない。しかしこの場合、私は正しさの規準を持っていない。ここで私にとっていつも正しく思われるものが正しい、と言いたくなるかもしれない。そしてそれは、ここでは「正しい」ということについては語りえないということではない。（PU-258）

記号「E」の適用を、記憶や想像に基づいて正当化するならば、記号の使用者は「正しさ」について語るができなくなる。ウィトゲンシュタインはそのような「主観的な正当化」（PU-265）に対して、「朝刊に書かれていることが本当であることを確かめるために同じ朝刊を何部も買う」（ibid.）、「想像の中で時計を眺める」（PU-266）、「橋の構造設計を正当化するために橋の材料の強度実験を想像する」（PU-267）、「『自分の身長くらい知っている』と言いながらその印として頭のとっぺんに手を置く」（PU-279）といった比喩表現を与えながら、それが正当化として「無意味」であることを示しているように見える。したがって、「正しさ」の規準が伴わないまま適用される記号は本当の意味での「言語」たりえない——多くの解釈者たちは、以上のような

考え方に基づいて、私的言語「E」の不可能性を主張するのである。

(3) 筆者の主張

第 258 節では、「E」が「感覚の記号」として導入される。この点に注意を払おう。① そもそも私的言語とは「他人には理解できない言語」であった。さて、「E」が私的言語であるならば、他人はこの記号の意味を理解できないはずである。だが、実際には、「E」は「感覚」という公共言語を伴いながら導入されるため、その意味は、他人にとって理解可能なものに成り下がっている。その限りで「E」は公共言語の圏域に属している (cf. 丸田 1998 p.80, 入不二 2006 p.110, 尾形 2010 p.192)。換言すれば「E」は感覚の文法に従うのである。そして次の命題からも伺えるように、感覚が私的であることも文法的に規定される事柄である。「『感覚は私的である』という文は『ひとはペイシェンスを一人でやる』という文に匹敵する」(PU-248)。以上を踏まえると、「E」という感覚の記号は、私的なあり方を文法的に保持したまま公共言語の中に位置付けられることになる。この記号は、私的な感覚を文法的に温存するからこそ、私的言語という当初の想定には到達できなくなる。② 次に規準の不在について。第 243 節が想定する意味での私的言語には、公共的に観察可能な規準が定義上存在しない。そうすると、規準の不在を根拠とする私的言語批判は、「私的言語には規準が存在しない」という、私的言語の「定義」から導かれる内容を単に繰り返して主張することによって、私的言語の不可能性を示そうとしているということになる。しかしながらそのような私的言語批判は、初めから「循環」している(「論点先取」を犯している)ように思われる (cf. 鬼界 2003 pp.316-317)。

①と②を踏まえると、「E」の適用に規準が伴わないからといって私的言語の不可能性が示されるわけではない、ということが分かる。そこで本稿は次のような可能性に着目したい。それは、「E」を規準なしに適用することが言語ゲームにおける新たな「原事実 (Urtatsache)」を成立させるような可能性である。そのような可能性は『探究』第 270 節において提示されている。

記号「E」を私の日記に書き込むことの一つの適用を考えよう。私は次のような経験をする。ある感覚があるときはいつでも、血圧計が私の血圧の上昇を示す。そこで私は、自分の血圧の上昇を計器の助けを借りずに言えるようになる。これは役立つ成果だ。私とその感覚を正しくそれと認めたかどうかは、この場合まったくどうでもいい。(中略)

どんな根拠があって、私たちは「E」をある感覚の名前と看做すのか? もしかするとこの記号がこの言語ゲームで使われている仕方が、その根拠な

のかもしれない。――では何故、「ある感覚」なのか。つまり何故、毎回同じ感覚なのか。それは毎回「E」と書くと仮定しているからだ。(PU-270) この箇所では、「E」を感覚の名前と看做して良いことになっている。だがその根拠は何か。一見すると、血圧の上昇は、「E」の適用の規準になっているように見える。とはいえ、「E」が感覚の名前であることが血圧の上昇という規準によって保証されるとすれば、血圧の上昇が発見される前には、この記号は無意味であったことになる。しかしながら、血圧の上昇の発見によって言えるようになるのは、あくまでも「E」の記入と血圧の上昇との相関関係に過ぎないのだから、この記号は、血圧の上昇によって有意味になるわけではない(相関関係は因果関係を含意しない)。この記号は、規準とは無関係に感覚の名前と看做されるべきであるⁱⁱⁱ。実際ウィトゲンシュタインは、血圧の上昇を、「E」が感覚の名前であることの根拠とは看做していない。むしろ彼が着目する根拠は、「E」が言語ゲームの中で恒常的に使用される仕方である。それゆえ、第 270 節で示唆されているのは、「正しさ」の規準なしに「E」を日記に書き込むことは批判されるべき難点ではない、ということである。語の無根拠な使用について、ウィトゲンシュタインは別の箇所でこう書いている。「ある語を正当化せずに (ohne Rechtfertigung) 使うということは、その語を不当に用いるということではない」(PU-289)。

「感覚日記」の事例では、記号「E」が私的言語の候補として導入された。にもかかわらず、(1)「E」は、「感覚」という公共言語へと読み換えられる。それどころか、(2) この記号を「正しさ」の規準なしに使用することは、感覚の言語ゲームにおける「原事実」へと繋がっている。従来の解釈は、以上の可能性を十分に捉えていないように思われる。他方、ウィトゲンシュタイン自身は、「感覚日記」の事例を通じて「対象と表記」(cf. PU-293) の関係を断ち切ろうとしているように見える。そこで本稿が取り上げたいのは、両者を結び合わせている「指示」という言語モデルである。

通常、「指示」や「命名」は、言語を構成する「制度 (Institution)」、「慣習 (Gepflogenheit)」に属している。したがって制度・慣習から逸脱して対象(例えば「感覚」)を指示しようと注意を集中したり、名(例えば「E」)を繰り返し唱えても、そのような行為は無意味なパロディーと看做されてしまう(cf. 鬼界 2003 p.318)。実際ウィトゲンシュタインは、指示の空転を示すために、多様な比喻を案出している。感覚に対する名の指示は、「右手から左手への現金贈与」(PU-268に喩えられており、そうした行為は謂わば「針に文字盤を固定して一緒に回るようにする」(BB p.71こと、「車での急ぎの旅行のときに、あたかも車を内側から押せるかのように、本能的に自分の前に

あるものを押す」(BB p.71)ようなこととして特徴付けられている。確かに第 258 節でも示唆されているように、「E」の指示は、結果的に無意味な「儀式 (Zeremonie)」と看做される。だが、本稿の考えでは、それが無意味な儀式と看做されるのは、「E」の指示が制度・慣習から逸脱しているからではない(第 243 節と第 258 節を注視すれば分かるように、私的な直示的定義はその定義からして制度・慣習の外部にあらざるをえない)。むしろ、記号の使用者自身でさえも記号が指示しようとしている対象(「感覚」)を同定しえないからである。以下では、「指示」や「同定」の内実を明らかにするために、『探究』への先行研究」と銘打たれている『青色本』の議論を援用する。

『青色本』では二つの異なる「私」の用法が設けられている。「客体としての用法 (the use as object)」と「主体としての用法 (the use as subject)」である(以下「客体用法」「主体用法」と略記)。この区別は内観の対象知覚的解釈に対する批判に基づいている。内観の対象知覚的解釈とは、自分の身体感覚を、外界の対象の知覚に見立てるような考えである。ウィトゲンシュタインによると、この解釈が否定される文脈において、上記の文法的な区別が設けられる。客体用法の場合、外界は必ずしも現れる通りには知覚されないため、誤認・誤同定の可能性が残される。ウィトゲンシュタインは「私の額にこぶがある (I have a bump on my forehead)」を例に挙げて次のように論じる。「(中略) 私は鏡を見て、隣人の額のこぶを私の額のこぶだと間違えることもありうる」(BB p.67)。「私の額にこぶがある」の「私」は特定の人物を指示するため固有名による置き換えが可能である。これに対して、主体用法の場合には、身体感覚は主体に現れる通りに知覚されるため、誤認・誤同定の可能性も生じない。例えば、「私は歯が痛い (I have toothache)」と発言する場合には、ある特定の人物を識別するという問題は生じない。『痛みを感じているのがあなただということは確かなのか』と聞くとすれば、それは無意味であろう」(BB p.67)。主体用法の「私」を固有名に置き換えることはできない。にもかかわらずウィトゲンシュタインは、他人がその「私」を特定の人物として理解しても良いことを自身の視点を交えながら論じている。

私は「私」という語によって L.W.を意味していない。しかし私はいま事実として L.W.である以上、他人が、「私」とは L.W.を意味するのだと理解しても、それで用は足りる (will do)。(中略) 私が望むのは、他人が私の言うことを理解するのは論理的に不可能であるということである。私の言うことを他人が理解すると語ることは、偽ではなく、無意味であるべきなのだ。(BB p.64 強調引用者)

本稿の見立てでは、『探究』の「私的言語論」の原型はこの箇所にある。そも

そも私的言語とは、「他人には論理的に理解できない言語」であった。そしてこれは、語の指示対象が同定されないことに起因している。つまり主体用法の「私」は、「E」と同様、私的言語の候補として導入されている。だからこそ、私的言語の候補（「E」や「私」）が挙げられた瞬間、それらの語は公共言語（「感覚」や「L.W.」）へと不可避的に読み換えられるのである。「私的言語論」の眼目はこの矛盾した状況の提示にこそある。それゆえ、「感覚日記」の事例が私的言語をなおも想定しようとする、「E」から「感覚」という規定をも取り除かなければならない。とはいえ、「感覚」を別の語に置き換えたとしても、それもまた公共言語に属するのだから、実のところ、私的言語という想定には、いつまで経っても到達できないのである。ウィトゲンシュタインは、『探究』第 261 節において、そのことを示唆している。

「E」を感覚の記号と呼ぶことに、われわれはどんな根拠を持っているのだろうか。何しろ「感覚」はわれわれに共通の言語の語であって、私だけに理解できる言語の語ではないのだから。したがって、この語を使うには誰もが理解するような正当化が必要である。——それが感覚である必要はない、彼が「E」と記すときには何かを持っている、という言い方をしてもまったく役に立たない。——しかも、それ以上のことは何も言えないだろう。けれども、持っている (haben) や何か (etwas) もまた、共通の言語に属している。——そういうわけで、哲学をしていると結局は、分節化されていない音声しか発したくないような地点に到達してしまうのだ。——だが、そういう音声表現であるのは、何らかの言語ゲームにおいてだけである。いまや、そうした言語ゲームが記述されるべきである。(PU-261)

(4) 今後の展望

これまでの内容を総括すると次のようになる。「E」と主体用法の「私」は私的言語の候補に挙がっていた。にもかかわらず、(1) それらは公共言語へと不可避的に読み換えられてしまう。つまり、「E」が「感覚」の名前として公共言語の中に位置付けられるのと同様に、「私」という語は、「L.W.」のような固有名として理解されてしまう。だが他方では、(2) 「E」や「私」の指示対象が同定されないため、「E」や「私」を、「感覚」や「L.W.」に置き換えることはできなくなる。(1)と(2)は矛盾している。この矛盾をどう理解すればよいのか。この矛盾から「語りえず示される独我論」を擁護する方法は導かれるのだろうか。それとも、語ることは失敗しても、まさにその失敗によって示されるものがある、という見解を否定する「決然たる読み手たち (resolute readers)」の考えにならって、上記の矛盾を「明白なるナンセンス」(cf. PU-

464)と看做すべきだろうか。筆者の「仮説」は次の通りである。『青色本』における「私」の矛盾と『探究』における「E」の矛盾とは、その形式や構造においては一致しているものの、中身や背景を異にしている。「私」の矛盾は独我論を語りえないものとして示しはするが、「E」の矛盾はそうではない。後者の矛盾は、あくまでも純然たる「ナンセンス (Unsinn)」という観点のもとで理解されるべきである。今後の研究課題は、そのことを『探究』の「治療哲学観」に照らして論証することにある。その際に手引きとするのは、S・カヴェルとS・マルホールの解釈 (Cavell 1979, Mulhall 2007) である^{iv}。

ⁱ 例えば、平田 2013 は、感覚表現の「一人称権威」を裏書きするものとして「不可謬性」と「私秘性」を挙げている (ibid. p.88)。

ⁱⁱ この概念は「徴候 (Symptom, symptom)」と対置される。例えば、私の喉が炎症を起こしていることは、私が嘔喉炎に罹っていることの「徴候」であるのに対して、私の血液中にしかじかの細菌が存在することは、私が嘔喉炎に罹っていることの「規準」である。前者は経験的なもの、後者は定義的なものである (cf. BB pp.24-25)。

ⁱⁱⁱ 血圧の上昇を規準ではなく徴候と看做す余地はある (cf. 丸田 1998 p.80)。

^{iv} *The Claim of Reason*, “Wittgenstein’s Private Language”を参照する。

(慶應義塾大学)

(5) 参考文献

- Bertolet, R, 1995, “Private Language Argument”, *The Cambridge Dictionary of Philosophy*, Robert Audi ed., Cambridge University Press.
- Glock, H, 1996, “Private Language Argument”, *Wittgenstein Dictionary*, Blackwell.
- Hacker, P.M.S., 1990, *Wittgenstein Meaning and Mind*, Basil Blackwell.
- Kripke, S.A., 1982, *Wittgenstein on Rules and Private Language*, Harvard University Press.
- Malcolm, N, 1954, “Wittgenstein’s Philosophical Investigations”, *Wittgenstein The Philosophical Investigations*, George Pitcher ed., Palgrave Macmillan, 1966.
- Wittgenstein, L, 1958, *The Blue and Brown Books*, Basil Blackwell.
- Wittgenstein, L, 2009, *Philosophische Untersuchungen*, Wiley-Blackwell.
- 入不二基義, 2006, 『ウィトゲンシュタイン「私」は消去できるか』, NHK 出版
- 尾形まり花, 2010, 「基準論は私的言語の不可能性を主張できるか」 『哲学』 No. 61

鬼界彰夫, 2003, 『ウィトゲンシュタインはこう考えた 哲学的思考の全軌跡 1912-1951』講談社現代新書

平田仁胤, 2013, 『ウィトゲンシュタインと教育——言語ゲームにおける生成と変容のダイナミズム——』大学教育出版

丸田健, 1998, 「『哲学探究』, 感覚日記の議論について」科学基礎論研究, Vol.25 No.2

世界3の自律性は言語的実体と衝突するのか

Does the Autonomy of World 3 Conflict with Linguistic Entities?

池田健人

Abstract

The world 3 theory, which Sir Karl Popper put forward in his later years, provoked a great storm of controversy. The most heated criticism has, among other things, focused on his argument for the autonomy of world 3. In this paper, I will first review the dispute concerning the reducibility of world 3 to the other two worlds and the conception of autonomy. Secondly, I will introduce the ontological analysis of the inmates of the three worlds.

Finally, I will suggest the importance of language for world 3.

(1) 研究テーマ

世界3の自律性と世界3における言語的実体の措定はポパーの三世界論においてどのようにして共に維持されるのだろうか。このことについて存在論の観点から考えていくため、その第一歩として本稿では言語の世界を新たに想定する議論がポパーの世界3概念に齟齬を来すということを明らかにする。

(2) 研究の背景・先行研究

不明瞭な直観としてではあるが長いあいだポパー哲学を根底から支え続けていた世界3概念は、その晩年、三世界論として公表された。三世界論とは、物理的対象ないし状態の世界（世界1）、意識の状態または心的状態、行動性向の世界（世界2）、そして思考の客観的内容の世界（世界3）という少なくとも3つの部分世界が存在論的に区別されうるとする議論である。これら3つの世界は、世界2の仲介によって相互作用することができる。

とりわけ、ポパー哲学の本質をなす世界3概念に対してはじつに多くの批判がある（Niiniluoto 2006: 59; Boyd 2016: 222, 230）。そのなかでも、もっとも厳しい批判にさらされている問題のひとつが世界3の自律性(autonomy)にかんするポパーの主張である（Gilroy 1985: 197）。

たとえば、ガデンヌ（Gadenne 2016）は物的一元論の立場から世界3の自律性を批判した。なぜなら、世界3が世界1に還元されうらば、その自律性は否定されるからだ。ポパーは、文明の再建における書物の重要性を指摘することによって、自律的な世界3の存在を示唆した（Popper 1972:

107-8)。しかし、ガデンヌによると、書物が世界に対して非常に大きな影響力をもつという帰結こそ妥当なものであるが、その事実を根拠として自律的な世界3の存在を主張することはできない。なぜなら、そのことによって私たちが結論できることは、たんに書物はある特殊な状況において世界の進行にかんして決定的な役割を演じることができるという事実のみだからである。ガデンヌによると、書物は明らかに世界1に属していて、結局のところ、私たちの世界は物理的なものによって著しく改変させられる。

また、世界3を世界2に還元しようとする議論もあるⁱ。パルスニコバ (Parusniková 2016) は認識主体の重要性を強調することによって世界3の自律性を批判した。問題とは私たちによって見つけられることを待っているようなものではなく、私たちの好奇心によって生命にもたらされるものである。いかなる認識主体ないし把握する精神も存在しないならば、同様にいかなる理論や問題もまた存在しないだろう。書物それ自体が背後に隠された可能的な理論や問題などを保有しているのではなく、それらの可能性は読者の可能性に依存している。

こうした批判に対して、どのように応答することができるだろうか。まず、物理世界は書物などのかたちで理論を取り込んだ物的なものによって変更されるというガデンヌの主張に対しては、その変化を生み出したのは書物の物的な側面ではなく、書物が理論それ自体を負っているという事実であるということで反論ができる。では、物理世界は理論それ自体ではなく、それらについての私たちの理解によって変更されるというパルスニコバの主張についてはどうだろうか。ポパーはこの種類の批判を重要であると認めている。というのも、世界1と世界3は媒介者としての世界2を通じてのみ相互作用が可能であり、理論それ自体による物理世界の変更にとって世界2の関与は不可欠だからである (Popper 1972: 148-9, 155; 1992: 184-5)。それゆえ、本節の残りでは、この問題について詳しく論じよう。

ポパーによると、世界3の対象は思考過程の内部において捉えられたものである (Popper 1972: 298-9)。ここで、そのことをわかりやすくするために、世界3の対象を可知的対象、そしてそれを捉えるまでの筋道を心的過程と呼ぶことにしよう。そうすると、ポパーは、思考過程の内部において可知的対象と心的過程を区別していたということができる。しかし、たとえそのような区別が導入されるとしても、それらはどちらも思考過程であるという点において違いはないのだから、パルスニコバにとっては、そこにはたんに2つの異なる世界2の対象があるだけのことにすぎないかもしれない。そこで、次のような例について考えてみよう。たとえば、いま私がよく知ってい

ある絵画について思い浮かべるとき、その思考過程、それゆえ可知的対象と心的過程の両者は明らかに精神世界に属するとパルスニコバは主張する。たしかに、パルスニコバのいうように、心的過程の結果として思い浮かべられた絵画、すなわち可知的対象もまた精神世界に属するものの一部であるということに変わりはない (Popper 1992: 181-2)。ところが、ポパーによると、心的過程の結果としての可知的対象は、精神世界に属する思考過程の一部であるにもかかわらず、それでも世界 3 に分類されなければならない。なぜなら、それは心的過程とは違い、さまざまな面からの批判的検討が可能なものだからである (Popper 1992: 181-2)。可知的対象は心的なものでありながら、同時に客観的なものでもある。パルスニコバが主張するように、これら両方が同じように精神世界に分類されるのならば、この点について明瞭な説明を与えることができないⁱⁱ。

このように、ポパーにとって、心的過程とその結果としての可知的対象は区別されるべきものであるということがわかった。それでは、次に、ポパーはいったいどのような意味で思考過程のうちでもとくに可知的対象に自律的な地位を与えたのだろうか。以下では、そのことを考えるために、世界 3 のさらに細分化された概念である世界 3.1、世界 3.2、世界 3.3 についてみていこう。

世界 3.1 とは、世界 1 のかたちで具体化ないし保存された世界 3 の部分である。たとえば、書物や芸術作品はもちろん、図書館や、人間の脳において記憶を司る領野などがこの世界に含まれる。世界 3.2 とは、いくらかの人間によって把握ないし理解された世界 3 の部分である。そして、爾余の領域が世界 3.3 ということになる。世界 3.3 とは、世界 1 として現実化されておらず、また世界 2 によっても意識的には捉えられていないような理論や問題ⁱⁱⁱ、あるいは将来的解決までをも含む世界 3 の部分のことを指す。この世界の存在者^{iv}は、以前にはいかなる主観によっても知られていることはない。またひょっとすると、それらは永遠に認識されることもないかもしれない。ポパーは、この世界をとくに影世界 (shadow world) と呼ぶ (Popper 1974: 1050-2; Popper & Eccles 1977: 41-3)。この世界 3.3 が、世界 3 の自律性を考えるうえで重要となる。

幾何学について考えてみよう。たとえば、たんなる直線や円、直角などは人間精神によって考案されたものである。しかし、たんなる直線や円、直角のみがあることと、直径に対する円周角が直角になること (タレスの定理) は同じではない。前者は発明されるものであり、後者は発見されるものである。幾何模様が私たちによって発明されたとき、タレスの定理もまた同時に

誕生した。しかしそのとき、人間精神によって捉えられ、世界 3.1 と世界 3.2 の両方に属することができたのは、あくまで直線や円、直角などの個別的な幾何模様のみである。タレスの定理は、依然として世界 3.3 に存在していて、のちにタレスによって—おそらくはタレス以前の誰かによって—発見されることとなった (Popper 1994: 26)。世界 3 はひとたび発明されると、その瞬間からひとり立ちをして、独自の生を歩み始める。

以上により、ポパーが世界 3 の自律性ということは何を意味していたのかは明らかとなった。ポパーは、私たちの制御を超えて世界 3 に潜むような対象、すなわち意図されざる結果 (世界 3.3 の存在者) を指して自律的な産物と呼び、またそのような対象が生じうる世界 3 の性質を自律性という言葉によって表現した。

しかし、ギルロイ (Gilroy 1985) によれば、ポパーによるこのような世界 3 の自律性の主張は、そこに必然的に含まれる不合理な帰結によって深刻に損なわれる (cf. Cohen 1980: 177; O'Hear 1980: 196–7)。たとえば、世界 3 がポパーのいうようなものであるとするならば、私たちが入手しうるあらゆる知識は、人間精神によって最初の矛盾が産出されたとき、自律的に派生したということになるのだろうか。なぜなら、ポパーは自己矛盾を抱えた理論もまた世界 3 に属すると主張するとともに、その真偽にかかわらず、あらゆる命題は矛盾から得られるということをも認めるからである (Popper 1963: 317–9; 1972: 126, 297; Popper & Eccles 1977: 56–7)。ほかにも、具体的に言及される論理法則には一貫性や含意、両立性、否定などを挙げるができる (Popper 1963: 298; 1972: 297–9; 1992: 180; Gilroy 1985: 187, 199)。ポパーの世界 3 において、ある命題から演繹ないし推論されるすべての論理的帰結は、その命題が生成されるやいなや自律的に生じる。ところが、このような事態は常識に反するのではないだろうか。

このような批判に対して、ニーニルオト (Niiniluoto 2006) は、ギルロイの議論においては自律的な世界 3 の構造が正しく考察されていないと指摘する。ある完全な言語 \mathcal{L} におけるあらゆる文の集合を仮定しよう。ここで、その言語によって定式化されるある一貫性のない理論を導入すると、そのとき、その理論の論理的帰結の集合はその言語において構成されるすべての文の集合と等しい。世界 3 をたんに可能的な文の集合であると考えれば、あらゆる命題はすでに世界 3 にあったのであり、その意味で世界 3 には何も新しいものは加えられないといえることができる。このように、ニーニルオトはギルロイの見解を認めつつも、ある命題が断言されないかぎり、非統一的な世界 3 における命題の集合が問題を引き起こすことはないことを主張する。なぜなら、

知識一般は断定された命題の全体、提出はされたがまだ反証のされていない命題の部類、あるいは実験を経て裏づけのされた命題などとして定義することができるからである。こうすると、私たちには限られた数の命題しか手に入らない。それゆえ、私たちが入手しうるあらゆる命題は、言語によって定式化されうる可能的対象としては世界3に存在しているが、人間精神によって最初の矛盾が産出されたときにすでに知識として存在していたということにはならない。

ここまでの流れを要約しておこう。世界3を世界1に還元しようとする議論に欠けているのは物理的対象のかたちで具体化ないし保存されている可知的対象への視点であり、世界3を世界2に還元しようとする議論に欠けているのは心的過程と可知的対象の峻別であった。しかし、可知的対象もまた思考過程の一部であることに変わりはない。それゆえ、可知的対象も世界2に属するのではないかという批判があるが、可知的対象は、心的なものでありながら同時に客観的なものでもあるという点において、たんなる心的過程とは一線を画するものであるため、自律的なものとして世界3に属すると考えられなければならない。ポパーによると、世界3の自律性を考えるうえで、世界3.3の概念は重要である。世界3.3の存在者は、私たちの発明の意図されざる結果として自律的に産出される。そのような意味で、ポパーは世界3の自律性を主張した。しかし、もしそうだとすると、私たちが入手しうるあらゆる知識は世界3においてずっと以前からすでに存在していたということになるのではないだろうか。たとえば、プトレマイオスの宇宙論が提出されたとき、その否定としてアインシュタインの宇宙論は当時からすでに存在していたということになるのではないか。ところが、ニーニルオトによると、世界3とはある言語における可能的なあらゆる文の集合であるという意味で自律的であるにすぎない。そのなかでもとくに私たちによって確かめられた命題のことを知識と定義するならば、非統一的な世界3が問題となることはない。

(3) 筆者の主張

ところで、ニーニルオトが世界3の自律性にかんするポパーの議論を言語の観点から擁護したのはなぜだろうか。その理由のひとつとして、ポパーが世界3の考察において言語の重要性を強調していたということが挙げられる。ポパーによると、世界3に属するあらゆる命題や理論、議論などはすべて言語によって定式化されうる。それゆえ、それらはたとえ誰によっても産出されなかったり、理解されなかったりしたとしても、それでも依然として誰か

によって把握されうるような可知的対象として世界3に属することができる (Niiniluoto 2006: 65; Popper 1972: 116, 137)。これは、言語が世界3に属しているということを意味している。しかし、この点について、ポパーの三世界論には概念的な混乱があるのではないかと指摘されることがある。

たとえば、クレムケ (Klemke 1979) によると、世界3とその存在者にかんする概念は調和しない。ポパーにしたがえば、世界3は論理内容の世界であるにもかかわらず、それと同時に言語によって定式化された理論や推測、問題、議論、書物などの世界でもまたある (Popper 1972: 74, 107)。それゆえ、ポパーの説明にしたがうと、たとえば書物について考えるとき、私たちは書物の論理内容と書物それ自体はどちらも世界3に属すると考えなければならない。

しかし、書物の論理内容は書物それ自体と同じ水準において分類されたり同定されたりすることができるようなものではない。たとえば、人類が滅亡したあとの状況を想像してみよう。それでも、ある文明化された私たちの後継者によっていくつかの書物が発見されるならば、それらは解読されることだろう (Popper 1972: 116)。ところで、このように書物が把握されたり理解されたりするというのは、その書物の文章によって伝達されている論理内容が解釈されているということにほかならない。したがって、書物が解読されるのはその論理内容によるのであって、それを表現する手段としての言語や書物それ自体によるのではない。それにもかかわらず、ポパーは世界1と世界3の両方に属するようなものとして、たとえば書物のような対象の存在を示唆した (Popper 1974: 1050-2; 1977: 41-3)。それゆえ、そのようなポパーの主張には問題がある。

そこで、クレムケは言語的要素 (linguistic component) に言及することによって世界4を導入する。クレムケによると、世界1には特殊領域があり、物理的対象のなかでも言語的要素を含んでいるものがこの領域の存在者であるとみなされる。それらは解釈される傾向性を有しているがゆえに認識内容を表現したり、世界1における他の存在者とのあいだに指示関係を構築したりすることができる。それゆえ、世界3に属することができるのは人間精神の産物ではない純粋な論理内容のみであり、その論理内容がひとたび言語によって定式化されるならば、したがってたとえばたんなる言語的実体としての理論や議論、ひいては物理的な書物などのかたちをとって具現化されるならば、それは世界4の存在者であるとみなされることとなる。

たしかに、言語ないし言語的要素を含むあらゆる存在者を世界4に分類するクレムケの指摘はもっともなものであるようにみえる。しかし、本当にそ

のような混乱がポパーの三世界論にはあったのだろうか。ポパーは「もろもろの理論、命題、言明は、もっとも重要な世界3の言語的実体である」(Popper 1972: 157)と述べ、言語は、世界1と世界2のみならず、世界3にも属するということを明言している。それどころか、ポパーは世界3について「言語の世界 (world of language)」や「言語的世界3 (linguistic third world)」などのように、あえて言語の世界であることを強調するような表現を用いることさえある (Popper 1972: 118, 120, 148) vii。つまり、クレムケの主張とは異なり、概念的な混乱があったからではなく、明確な意図によってポパーは言語を世界3の対象とみなしているのである。そうすると、ポパーの三世界論においては、言語はむしろ世界3の存在者でなければならない。

(4) 今後の展望

ポパーは問題や理論、批判的議論の世界を世界3の特殊ケース、狭義の世界3、世界3の論理的または知的領域であるとみなしている (Popper 1992: 187)。このことから、クレムケが世界3を論理内容の世界であると結論したのに対して、ポパーはあくまで論理内容は世界3の一部にすぎないと考えていたということがわかる。しかし、そこでポパーはなぜ言語を世界3の存在者に加えたのだろうか。たとえ論理内容は世界3の一部にすぎないとしても、それはすなわち言語が世界3に分類されるということを意味しているわけではない。ポパーは、いかにして世界3の自律性と世界3における言語的実体の存在を調停したのだろうか。三世界論における言語の重要性はポパーが繰り返し主張しているところではあるものの、言語が世界3の対象としての地位を獲得するまでの経緯については改めて詳しく見直される必要がある。

i 認識主体なき認識論の不可能性については、Bunge (1981) においても、進化論的認識論の観点からの議論が確認できる。

ii ジェームズ (James 1981) のように、思考過程の内部における区別の存在は認めるものの、それでもなお可知的対象の自律的な世界については心理学的観点から否定しようとするような議論もある。

iii そのような理論や問題は、たとえ意識的には捕捉されていなくても、直観的には世界2によって関心が寄せられうる。しかし、それはあくまで私たちを自律的な世界3.3における問題の探求へと駆り立てる要因にすぎない (Popper 1974: 1052)。世界2によって世界3.3へと直観的な関心が向けられている状態は、世界2において世界3の対象が把握されたり理解されたりしている状態とは区別されるので、世界2によって直観されているのみの対象は依然として世界3.3に分類される。Popper (1977) p. 43 ff.も合わせて参照されたい。

iv 本稿では「存在者」という語によって統一するが、これは実際にポパーが用いている表現ではない。ポパーは世界3の対象を住人 (inmate) や市民

(citizen)、居住者 (inhabitant) などの語によって表現する。また、世界 3 に言明を住まわせる (people) というような表現も認められる。三世界論において「住まわせる」という語には少なくとも 2 種類のもので用いられており、筆者の知るかぎりでは、もう一方の住まわせる (inhabit) という語は、ポパーによってはアイデアや精神に対してのみ用いられるが、ポパー研究においては世界 3 の対象にも適用されることがあるので注意されたい (Popper 1972: 107, 123, 126, 138; 1992: 182; e.g. Gilroy 1985: 187)。

v ポパーによると、世界 3 における発見はその証明とは関係がない。ただ既存の概念などに対して直観的に疑念を抱いたり、あるいは違和感を覚えたりするのみで十分である (Popper 1994: 35)。

vi ここで意図される言語はよく定義された形式言語であり、自然言語によっては充足されえないある理想的状況が想定されている (Niiniluoto 2006: 64)。

vii ポパーは 1970 年ごろを境にエックレスにしたがって世界の呼び方を変更している (Popper 1972: vii; 1992: 181)。本稿では変更後の呼称を採用しているため、統一性のために third world は「世界 3」と訳している。

(5) 参考文献

Boyd, B. 2016. 'Popper's World 3: Origins, Progress, and Import', *Philosophy of the Social Science* 46(3): 221–41.

Bunge, M. 1981. *Scientific Materialism*, London, England: D. Reidel.

Cohen, J. 1980. 'Some comments on Third World Epistemology', *British Journal for the Philosophy of Science* 31(2): 175–80.

Gadenne, V. 2016. 'Is Popper's Third World Autonomous?', *Philosophy of the Social Science* 46(3): 288–303.

Gilroy, J. D. 1985. 'A Critique of Popper's World 3 Theory', *The Modern Schoolman* 62(3): 185–200.

James, W. 1981. *The Principles of Psychology, Vol. I*, Cambridge, Massachusetts and London, England: Harvard University Press.

Klemke, E. D. 1979. 'Karl Popper, Objective Knowledge, and the Third world', *Philosophia* 9(1): 45–62.

Niiniluoto, I. 2006. 'World3: A Critical Defence', *Karl Popper: A Centenary Assessment, Vol. II: Metaphysics and Epistemology*, ed. I. Jarvie, K. Melford, and D. Miller, pp. 59–69, Aldershot: Ashgate.

O'Hear, A. 1980. *Karl Popper*, London, Boston and Henley: Routledge & Kegan Paul.

Parusniková, Z. 2016. 'The Devaluation of the Subject in Popper's Theory of World 3', *Philosophy of the Social Science* 46(3): 304–17.

Popper, K. 1963. *Conjectures and Refutations: The Growth of Scientific*

Knowledge, New York: Basic Books. (『推測と反駁』、藤本隆志・石垣壽郎・森博訳、法政大学出版局、1980年)。

Popper, K. 1972. *Objective Knowledge: An Evolutionary Approach*, The University Press. (『客観的知識—進化論的アプローチ』、森博訳、木鐸社、1972年)。

Popper, K. 1974. 'Replies to My Critics', *The philosophy of Karl Popper*, ed. P. A. Schilpp, pp. 961–1197, La Salle, Illinois: The Open Court.

Popper, K. 1992. *Unended Quest: An Intellectual Autobiography*, London: Routledge. (『果てしなき探求—知的自伝』、森博訳、岩波書店、1978年)。

Popper, K. 1994. *Knowledge and the Body-Mind Problem: In Defence of Interaction*, ed. M. A. Notturmo, Routledge.

Popper, K. & Eccles, J. C. 1977. *The Self and Its Brain*, Springer International. (『自我と脳 (上・下)』、西脇与作・大村裕訳、思索社、1986年)。

(大阪大学)

予測符号化の哲学的含意と過学習の問題
Philosophical Implication of Predictive Coding and the Problem of
Overfitting

柴田 翔平

Abstract

The aim of this paper is to show the philosophical relevance of predictive coding, a theory of perception in neuroscience. In philosophy of perception, not much attention has been paid to predictive coding. However, it can elucidate a philosophical problem which has not been discussed much, namely the relationship between perception and its generality. I start by introducing predictive coding with a statistical method necessary for understanding it and then argue that we see not only individual objects immediately present in our perception, but also general aspects of the objects.

1 研究テーマ

本論文では神経科学のみならず哲学においても影響力を増しつつあるリサーチプログラムである予測符号化を用いて知覚の哲学への含意を検討する。予測符号化に従えば知覚は神経系において事前知識を用い、感覚入力の原因を予測する過程である。予測符号化はより一般化され「自由エネルギー原理」として知覚に限らず行為や認知、価値などの多岐に渡る現象の説明に用いられている [1]。哲学的な関心も高まっており、Clark[2] や Hohwy[3] が知覚の哲学に応用している。予測符号化の特徴として統計・機械学習の手法を応用していることが挙げられるが、しかしこの点に注目した哲学的研究はまだ少ない。そこで予測符号化のこの点の応用の一例として知覚における過学習の問題を取り上げ、そこから知覚は個物にのみ関わっているのではなく、一般性の度合いを持ちうることを示唆する。この結論は知覚が目前の個物に関するものであるという典型的な理解と対立することとなる。そうした典型的な理解として次の一節が代表的である [4]。

冬の最中に春の庭や、秋の葉っぱに覆われた庭を想像できるし、現在そうではないようなあらゆる仕方でその庭について考えることができる。これは知覚においては可能でない。なぜなら知覚は現在与えられているもののみ直面しうるからである。

本論文における過学習からの議論が成功していれば、知覚の内容は現在与えられているものに限定されておらず、一般性の度合いを持つことが示唆される。予測符号化の統計的な側面を哲学的議論に持ち込むという新たな観点

から、知覚対象が個物なのか一般者なのかという知覚の哲学における議論 [5] に貢献することが本論文の目的である¹。

本論文で示すのは以下のことである。(1) 知覚 (特に視覚) はベイズ推論によって捉えられること、(2) ベイズ推論は予測誤差最小化によって実現されていること、(3) 予測誤差最小化は統計的な手法であり過学習しうること、(4) したがって知覚でも過学習が起きうること、(5) 知覚における過学習は対象の持つ個別偶然的な性質に注意を過剰に向けることの帰結であることを示す。

知覚の哲学において予測符号化や統計的手法は頻繁に用いられないことを考慮し、数理的な説明は議論にとって必要最小限に留める。最小限の数理的な説明ののち、予測符号化の持つ哲学的な含意を検討する。

2 研究の背景・先行研究

予測符号化を理解するためには、知覚が一对多対応の関係を一对一对応にもたらず過程であることを認識する必要がある。つまり、私たちがなにかを見ているとき、見ているものは得られた感覚入力から一意的に得られるものではなく、知覚におけるなんらかの作用によって複数の候補の中から決定されているのである。というのも、感覚入力とその原因の間には一对多対応が成り立っているため、感覚入力だけでは可能な原因が複数存在してしまうからである。これは網膜が二次元であることを考えると分かりやすい。網膜上に映る像は二次元であるにも関わらず、私たちの見る対象はあたかも三次元上のものである。しかし、二次元上の対象を三次元に投射する仕方は一意的に決まらない。得られる感覚入力は二次元上のものであり、目の前の対象が自分からどれだけ離れているかという情報はそこにはないため、私たちが見ているものは単なる網膜上の像以上のものである。

このために、知覚は場合に応じて可能な候補から一つを選び取る仕組みを持っていると考えるのは自然であるが、この仕組みが何かというのは自明ではない。この仕組みを明らかにするのが知覚研究における課題の1つである。Rao and Ballard[6] は網膜における感覚入力が視覚野などの高次の領域に伝達されるだけでなく、高次の領域から低次の領域へと感覚入力に関する予測が伝達されることを発見した。予測符号化においては感覚入力と予測に基づき、ベイズ推論によって最も蓋然性の高い感覚入力の原因を推定し、その推定結果が知覚内容になると主張される。こうした観点から知覚を説明した研究として Buckley et al.[7] を解説する。

Buckley et al. はまず、周囲の状態 θ は直接知覚できず、感覚入力 φ から推測されなければならないということを仮定する。ただし、 $p(\theta)$ と $p(\varphi)$ は θ と φ が取りうる値の確率分布であり、 θ と φ は开区間 $(0, 1)$ に実数値を取る

確率変数である。例えば物体の凹凸の度合いが ϑ 、網膜における光量が φ で表されるとすると、それぞれの確率分布は最もよくある凹凸度合いや光量に高い確率を割り当てる。ここで感覚入力 φ から状態 ϑ の推測をしたい。それを条件付き確率によって表したのが、「ある感覚入力を与えられたときの状態に関する確率分布」である事後分布 $p(\vartheta|\varphi)$ である。そしてそれを計算すること、つまり感覚入力に基づいた状態の推測を可能にするのがベイズの定理である。ベイズの定理より以下の等式が成り立つ。

$$p(\vartheta|\varphi) = \frac{p(\vartheta)p(\varphi|\vartheta)}{p(\varphi)}$$

事後分布 $p(\vartheta|\varphi)$ はある感覚入力 φ のもとで、周囲の状態 ϑ が取りうる確率である。ここで注意すべきは、 φ がすでに観測された一つの値であるのに対して、 ϑ は実現しうるであろう値であって、 $p(\vartheta|\varphi)$ はそれがどのように分布しているかを確率的に表現しているということである。事前分布 $p(\vartheta)$ は周囲の状態 ϑ がどのように分布しているかを確率的に表している。エージェントは直接周囲の状態を知ることはできないが、これまでの経験からそれに関する信念を持つことはできるだろう。例えば、究極的には知り得ないことであるが、ある対象がどのような凹凸を持っているかということは、これまでのその対象との接触経験から設定することができる。尤度 $p(\varphi|\vartheta)$ の解釈は注意が必要である。尤度も事後分布と同様に条件付き確率の形式をしているが、一般にここでも変数となるのは状態 ϑ である。そのため尤度はどの状態を所与とするとある感覚入力 φ が尤もらしいかを表していると解釈できる。残りの周辺尤度 $p(\varphi)$ は感覚入力 φ に関する確率分布であるが以下の式で求められる。

$$p(\varphi) = \int d\vartheta p(\vartheta, \varphi) = \int d\vartheta p(\vartheta)p(\varphi|\vartheta)$$

そのため事前分布と尤度の情報さえあれば事後分布を求めることができる。従ってベイズの定理においては ϑ のみを変数であり感覚入力 φ は固定された値であるから、ベイズ推論の使用は与えられた感覚入力 φ から状態 ϑ を予測するという目的に合致している。

しかし、周辺尤度を求めるための積分は計算機でさえもしばしば困難であるため、これを脳が行なっているというのはとても強い仮定である。そこで、 $q(\vartheta)$ という別の確率分布を考え、 $q(\vartheta)$ による $p(\vartheta|\varphi)$ の積分を用いない近似を考えることとする。 $q(\vartheta)$ は認識モデルと呼ばれる。Buckley et al. はここから $p(\vartheta|\varphi)$ と $q(\vartheta)$ のカルバック・ライブラー距離を求め、ベイズ自由エネルギーの最小化のためにラプラス近似を行う。途中の導出は省略するが、最終的に認識モデルによる事後分布の近似のために必要な最小化の対象は以下のようになる。

$$\begin{aligned}
E(\mu, \varphi) &= -\ln p(\varphi|\mu) - \ln p(\mu) \\
&= \frac{1}{2\sigma_z} \varepsilon_z^2 + \frac{1}{2\sigma_w} \varepsilon_w^2 + \frac{1}{2} \ln(\sigma_z \sigma_w)
\end{aligned}$$

ただし、 $\varepsilon_z \equiv \varphi - g(\mu; \psi)$ 、 $\varepsilon_w \equiv \mu - \bar{\mu}$ とする。ここで μ は ϑ の確率分布を表現するための脳状態を表す確率変数であり、 $\bar{\mu}$ はその平均、そして $g(\mu; \psi)$ は μ を変数、 ψ をパラメータとして取る関数であり、 φ の予測値である。すると ε_z は実際の感覚入力 φ と、その予測値である $g(\mu; \psi)$ の誤差を表しており、 ε_w は実際の脳状態 μ とその平均の $\bar{\mu}$ の間の誤差を表していると解釈できる。そして、それぞれ φ の分散 σ_z と μ の分散 σ_w の逆数によって重みづけられている。最後の項は変数 φ と μ を含まない定数であるから最小化に当たって無視できる。すると、分散の逆数によって重みづけられた感覚入力の予測誤差と脳状態に関する予測誤差の最小化問題に帰結する。そのため感覚入力から状態の推測は予測誤差最小化によって達成される。予測誤差最小化については過学習の問題を論じる際に再び触れることとする。

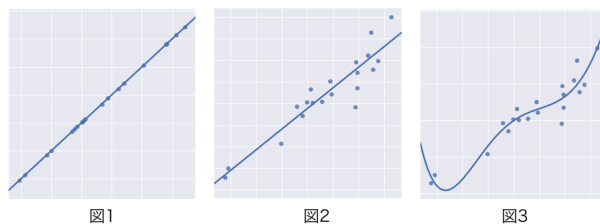
ここまでの Buckley et al. によるベイズ推論の近似に関する説明である。しかしなぜ他の近似法ではなくラプラス近似を使うのかに関する説明がなかった。ベイズ推論は様々な近似法がある。しかし、予測符号化理論の発端である Rao and Ballard の発見は感覚入力と予測が神経系における高次の領域と低次の領域の間で伝達されることだけでなく、2つの間の誤差が最小化されるということも含んでいる。この研究成果と整合的である理論を作ることが知覚研究における課題である。そのため知覚をベイズ推論によって特徴づけるためには神経系において予測誤差最小化が起きているという経験的制約と整合的であることが求められる。ラプラス近似の使用はこの観点から正当化できるだろう。

このようにして、予測誤差最小化が導かれた。ここから (1) 感覚入力の予測値と実際の値、そして (2) 脳状態の予測値と実際の値の誤差を 0 に近づけることが要請される。そして、感覚入力と脳状態のそれぞれの誤差は分散の逆数によって重みづけられる。例えば、感覚入力の誤差が小さく、脳状態の誤差が大きかったとしても、感覚入力誤差の分散が小さければ、結果として誤差は大きくなり、脳状態の誤差を上回ることもあり得る。その場合、脳状態の誤差よりも感覚入力の誤差を小さくするように動機づけられるだろう。この点はのちに取り上げる過学習の問題の際に重要になる。なお、Rao and Ballard に従えば予測誤差最小化は神経系のそれぞれの領域で行われるため、Buckley et al. はこの後予測誤差最小化を階層構造に拡張している。しかし、ここの

目的には一つの層の中での予測誤差最小化で十分であるためこれ以上は踏み込まない。

3 筆者の主張

これまでこうした統計的な側面に注意を払った予測符号化の哲学における応用はされてこなかったが、こうした側面にも哲学的問題がある。その一例として過学習の問題を考える。過学習とは統計的学習理論の用語であり、統計モデルがすでに持っているデータに対して過度に適合してしまうために予測精度が低くなってしまふことを指す [8]。特にデータの分散が大きく不確実性が大きい時に、強引にパターンを見出そうとするとそのデータにとって偶然的な性質までもパターンとしてみなされてしまう。なおここでは不確実であるということ、分散が大きいことと交換可能なものとして扱う。つまり、何かの不確実であると言う時の主語は一つではなく、複数の感覚入力や脳状態などのデータである。データの分散はデータ全体の平均からの離れ度合いとして定義されるから、分散が大きいというのは予測が難しく不確実な状態ということである。逆にデータの分散が小さく平均の周囲に密集していれば予測は簡単だろう。例えば人間の身長と体重に関するデータを得た場合、それは普通 1 次関数に従う過程によって得られるが、測定誤差を含んだデータに適合するパターンを作ろうとすればデータは直線から少しずれるため 1 次関数よりも複雑なパターンになる。しかしそうしたパターンは今後得られるデータを適切に予測できなくなるだろう。そのため誤差を無視する必要があるが、過度にデータに適合してしまうとこうしたことが起き、これが過学習として知られる現象である。



以上は (図 1) 分散の小さいデータでの予測誤差最小化と、(図 2) 分散の大きいデータでの 1 次関数による予測、(図 3) 分散の大きいデータでの 5 次関数による複雑な予測である。図 1 の場合には適切な予測を立てることは容易だが、図 2 と図 3 の場合には容易ではない。ここでは図 2 と図 3 は図 1 のデータにノイズを加えて発生させたものである。そのため適切な予測は図 1 と同

じ、つまり図2である。他方で図3はこのデータにおける偶然的なノイズに適合し過学習しているため、予測がうまくいかなくなるだろう。図3よりも適合の度合いが強い予測も考えることができるが、そうした予測が現在のデータに偶然的な性質に適合してしまっているのは明らかだろう。

知覚の話に戻すと、知覚は予測誤差最小化を行なっているのであったが、それは以下の式で特徴付けられるのであった。

$$E(\mu, \varphi) = \frac{1}{2\sigma_z} \varepsilon_z^2 + \frac{1}{2\sigma_w} \varepsilon_w^2 + \frac{1}{2} \ln(\sigma_z \sigma_w)$$

これはそれぞれの誤差に関する項にデータ（感覚入力、脳状態）の分散の逆数が乗じられている。データの分散はデータが平均から離れているほど大きくなる。その場合には分散の逆数は小さくなるため、誤差を小さくしなくても最小化したい値 $E(\mu, \varphi)$ は小さくなる。逆に予測誤差が小さくても分散が小さければ、 $E(\mu, \varphi)$ が大きくなってしまうため、予測誤差をさらに小さくする必要がある。データの分散が小さく不確実性が小さい時には細かいところまで予測誤差を小さくする必要がある。そのため予測と大きく異なる感覚入力を得た場合、それと適合するように予測を改訂する必要がある。一方データの不確実性が大きい時に同様の仕方では予測を改訂すると、図3で見たようにデータへの適合が過剰になり過学習が生じてしまう。

感覚入力や脳状態が不確実であればあるほど分散が大きくなり弱く重みづけられるようになるため、予測誤差を小さくする必要性が少なくなる。分散が大きい分布から得られたデータに対して過剰に誤差を小さくしようとする、そのデータの持つ偶然的な特徴にばかり適合してしまい、同じ確率分布から得られるであろう未知のデータに対する適合度が低くなってしまふ。不確実性の中に無理やりパターンを見出そうとすれば、今後得られるデータにはないであろう特徴を見出してしまふ。そして二次元上の像を三次元に投影する際に予測を使っているのであるから、予測が失敗すると正しい投影ができなくなり、視覚にも問題が出てしまふ。

どんな知覚も以前と同じ状況ということはあるから、知覚において予測をするためには目の前の対象を一般性において捉える必要がある。例えば目の前の机を見る時、その机の見え方は今まで見てきた机と完全に同一ということはないだろう。しかし、机が絵の中にある二次元上の対象ではないものとして、そして適切な距離感にあるものとして知覚するためにはこれまで机がどのように見えてきたかという情報を用いる必要がある。そのためにはこれまで見てきた机がどれだけ多様であったか（つまり不確実であるか）に応じて、予測誤差を小さくする必要がある。より不確実なときには目の前の机が持つ個別的な性質とこれまでの経験に基づく予測の間の誤差を小さくし

ざるべきではないが、これは知覚内容が必ずしも目の前の個別的对象ではなく、一般者にも関わりうることを含意する。予測誤差を小さくせずそのままにすると、知覚内容は目の前の机だけでなく、机一般にも関するものとなっている。仮に目の前の机の天板がこれまで見てきた机と微妙に異なる見た目であったとしても、予測誤差を小さくしようとして予測を改訂すれば、過学習の問題が示唆するようにその後の予測がうまくいかなるだろう。知覚が現在与えられているものに直ちに反応しなければならないとする知覚の典型的な理解はこの点を見落としている。得られる感覚入力の不確実な時には予測誤差を減らそうとしすぎれば予測誤差最小化によるベイズ推論の近似が成り立たなくなり、感覚入力から状態の推測がうまくいかなってしまう。

こうしたことが実際に知覚において起きうることの論拠として、自閉症スペクトラム患者の特徴的な知覚を過学習の結果とする研究が存在する [9]。そうした研究によれば、例えば自閉症スペクトラムの症状の一つである定型発達者ならば気づかないような対象の特徴に注意が向く傾向は、分散の逆数によって規定される度合いを超えた予測誤差の最小化によって過学習が起きると解釈される [10]。そうした例としてあげられるのがコミュニケーション時の顔の知覚である。人の顔は多様であり不確実であるため、目の位置や鼻の形、肌の質感などの予測誤差を下げようとしすぎると過学習してしまう。過学習すると感覚入力と知覚内容の対多関係の解決に必要な予測がうまくいなくなってしまう。顔を顔として認識できなければ相手の気持ちを顔から読み取ることも難しくなるためコミュニケーションが困難になるというのは自然だろう。そのためこれまでの経験に基づいた予測と異なる顔を見たとしても、今見ている顔に過剰に適合するような予測をするべきでないこと、そして個別的な顔を通して一般者としての顔を知覚する必要があることが以上の議論から理解できるだろう。

4 今後の展望

今後の展望としては、(1) ここで提示した知覚観を用いることで自閉症スペクトラムの特徴的な知覚をより詳細に分析すること、(2) 予測符号化をより一般化し、知覚のみならず行為や意識、価値などを説明することを試みる自由エネルギー原理においてここで検討した一般性の概念がどう関連しているかということ、(3) 予測符号化とは異なる観点から知覚の対象に一般性を帰する立場 [5] との関連性を検討すること、(4) 過学習の問題以外の予測符号化が用いている統計的手法が持つ哲学的含意を検討すること、これら 4 つの方向での研究を進めることが必要となるだろう。自由エネルギー原理において行為もまた予測誤差最小化を行なっているとされており [1]、現在は特に行為にお

ける過学習の研究を進めている。

注

¹ しかしながら既存の議論は予測符号化の研究を用いてこなかったため、ここでは予測符号化がそうした議論と関連しすることを示すに留め、既存の立場とどのような関係にあるかということは今後の研究の課題とする。

文献

- [1] Karl Friston. The free-energy principle: a unified brain theory? *Nature reviews neuroscience*, 11(2):127, 2010.
- [2] Andy Clark. *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford University Press, 2015.
- [3] Jakob Hohwy. *The predictive mind*. Oxford University Press, 2013.
- [4] Tim Crane and Craig French. The problem of perception. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 edition, 2017.
- [5] Hannah Ginsborg. Perception, generality, and reasons. In Andrew Reiser and Asbjørn Steglich-Petersen, editors, *Reasons for Belief*, pages 131–157. Cambridge University Press, 2011.
- [6] Rajesh PN Rao and Dana H Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature neuroscience*, 2(1):79, 1999.
- [7] Christopher L Buckley, Chang Sub Kim, Simon McGregor, and Anil K Seth. The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology*, 2017.
- [8] Douglas M Hawkins. The problem of overfitting. *Journal of chemical information and computer sciences*, 44(1):1–12, 2004.
- [9] Sander Van de Cruys, Lee de Wit, Kris Evers, Bart Boets, and Johan Wagemans. Weak priors versus overfitting of predictions in autism: Reply to pellicano and burr (tics, 2012). *i-Perception*, 4(2):95–97, 2013.
- [10] Sander Van de Cruys, Kris Evers, Ruth Van der Hallen, Lien Van Eylen, Bart Boets, Lee de Wit, and Johan Wagemans. Precise minds in uncertain worlds: Predictive coding in autism. *Psychological review*, 121(4):649, 2014.

(北海道大学)

新進研究者 Research Note

根元的解釈の基礎について
On The Basis of Radical Interpretation

中谷内悠

Abstract

In this paper, I consider Davidson's thought on a theory of interpretation of meaning, that is a theory of radical interpretation. One important criterion for a theory of interpretation to be good is that meaning is appropriately interpreted from the basis of interpretation. The other is that the basis of interpretation can be easily known based on observed facts by an interpreter. I show that the theory of interpretation presented by Davidson meets the first criterion, but does not meet the second criterion.

(1) 研究テーマ

この論文は、どのようにして意味（や命題的態度の内容）という概念に良い説明を与えられるか、という問題に関わる。例えば、ある未知のことばを話す外国人が、ある文を発したのを見て、私は同じ文を使って彼に聞き返したとする。もしそのとき彼がうなずけば、彼はその未知な文を真だとみなしていることがわかるだろう。つまり、行為者がある文を真とみなしているかどうかは、観察された事実をもとに、じかに知ることができる。しかし彼のふるまいを見て、その文が何を意味するのかということまで知ることはできない。このように、意味は観察可能な証拠からじかに知りうるものではないのである。この意味で、意味（や命題的態度の内容）という概念は、私たちにとって理解しづらい概念である。そして、意味（や命題的態度の内容）という概念を、観察可能な証拠からじかに適用可能な概念によって説明することは、それらの概念をより理解しやすい概念によって説明するという点で、良い説明である。D.デイヴィッドソンは、根元的解釈に関する考察によって、この種の説明を与えようとしている。彼は、意思決定理論を利用し、それを自身の意味の解釈理論と合わせることで、比較的理解しやすい行為者の態度のパターンをもとに、意味（や命題的態度の内容）といった比較的理解しづ

らいものへと至る道筋を明確に示している(Davidson,1980,1984)。この論文のテーマは、デイヴィドソンのこの試みが成功しているかどうかを検討することである。

(2) 研究の背景・先行研究

根元的解釈とは、ある行為者が使う未知の言語を、彼のふるまいをもとに解釈することである。根元的解釈の方法を明らかにすることは、意味を前提としないところから、行為者のふるまいと、そこから知ることができる態度をもとに、意味を解釈する方法を明らかにすることである。それゆえ、それは意味についての説明となっているのである。

根元的解釈について考えるにあたって、デイヴィドソンははじめ、ある文を真とみなすという話者の態度を解釈の基礎として、意味の解釈を行うことが可能だと考えた(Davidson,1973,1974)。しかし後に、この考えが適切ではないと考えるようになる(Davidson,1980)。まずはその理由について述べる。

意味を解釈するためには、話者がどの文を真とみなすかということがわかっているだけでは十分ではなく、話者がどの文をどの程度の度合いで真とみなしているのか、ということがわかっているなければならない。つまり、信念の度合いが必要となる。では、なぜ意味の解釈のために信念の度合いが必要となるのか。それを理解するためには観察文と理論的な文の関係性を理解しなければならない。例えば、「太郎の叫び声を聞いてイノシシは逃げた」のような観察文は、ひとつの観察された事実から真であることがわかるが、一方で「イノシシは大きな音を恐れる」のように、理論的な文は、ひとつの観察された事実から真だとわかるわけではなく、証拠となるその事実によって真であることがある程度で支持されるようなものである。そして、理論的な文の意味は、どの観察文によって、その理論的な文が真であることがどの程度支持されるのかということによって決まる。他方、観察的な文の意味は、その文が真であることによって、どの理論的な文が真であることがどの程度支持されるのか、ということによって決まる。それゆえ、文の意味を解釈するためには、話者がある文を真とみなすことによって、別の文が真であることがどの程度支持されると考えているかがわかっている必要がある。そして、この支持の関係は、「ある文にたいする信念の度合いが変化するのに伴って他の文にたいする信念の度合いがどのように変化するかに着目することによって」知ることができる(Davidson,1980,p155)。以上の理由で、意味の解釈のためには信念の度合いが必要となるのである。

しかし、信念の度合いは、観察された事実からじかにわかるわけでも、真とみなすという態度をもとに解釈可能なわけでもない。そういうわけで、ある文を真とみなすという態度をもとに、意味の解釈を行うことはできないとデイヴィドソンは考えるようになる。

そして、意味の解釈のためには信念の度合いを知ることが必要となるので、意味を前提としないで信念の度合いを解釈することが必要となる。デイヴィドソンはジェフリーの意思決定理論をもとにして、その方法を明らかにする。

意思決定理論は、行為者の選好をもとに、信念の度合いを導く方法を示している。例えば、ラムジーは賭けについての選好をもとに、それらを導く方法を明らかにする(Ramsey,1950)。一方で、ジェフリーの意思決定理論は、命題が真であることに対する選好をもとに、信念の度合いと望ましさを導く(Jeffrey,1983)。ここでひとつ注意しておきたいのは、賭けについての選好のランキングは、賭けの結果(1万円もらえること)の望ましさのランキングをそのまま示しているわけではないのに対し、命題が真となることへの選好のランキングはそのまま、命題の望ましさのランキングとなるということである。

まずは、命題が真となることに対する選好から、信念の度合いと望ましさを導く方法に関するジェフリーの考えを大まかに整理する。まず、ある命題の望ましさについてジェフリーは次のように考える(Jeffrey,1983,p.78-79)。例えば「行為者が2番の馬に1万円賭ける」という命題が真となるケースは、その命題が真となるのと同時に、例えば2番の馬が1等になるケースや、2等になるケースや、3等になるケースや、…など、可能なケースはいろいろとある。そして、「行為者が2番の馬に1万賭ける」という命題の望ましさは、この命題が真となるそれぞれのケースの望ましさを、それぞれのケースの確率で加重平均したものとなる。この考えは直観的に支持できるように思われる。そして次の、望ましさ公理はこの考えを反映するものである(prob(A)はAの確率を、des(A)はAの望ましさを表す)。

。

prob(XかつY)=0、かつprob(XまたはY)≠0ならば、

$$des(XまたはY) = \frac{prob(X)des(X) + prob(Y)des(Y)}{prob(X) + prob(Y)}$$

望ましさ公理が述べていることは、両立不可能な命題から成る選言の望まし

さは、 X が真となるケースや Y が真となるケース（これらは両立しない）の望ましさを加重平均であり、そのときの加重は、それらのケースの確率であるということである (Jeffrey, 1983, p. 80-81)。そして、 $Y = \bar{X}$ とすると、 $(\text{prob}(X) + \text{prob}(\bar{X}) = 1$ なので) 望ましさを公理から次の式が導かれる。

$$\text{des}(X \text{ または } \bar{X}) = \text{prob}(X)\text{des}(X) + \text{prob}(\bar{X})\text{des}(\bar{X})$$

そしてここから、次の式が得られる。

$$\text{prob}(X) = \frac{\text{des}(X \text{ または } \bar{X}) - \text{des}(\bar{X})}{\text{des}(X) - \text{des}(\bar{X})}$$

つまり、ある命題の主観的確率は、「論理的真理 (X または \bar{X}) の望ましさとその命題の否定の望ましさの間隔」と、「その命題の望ましさとその命題の否定の望ましさの間隔」の比となる (Jeffrey, 1983, p. 114-115)。

さて、選好のパターンから、これらの間隔の比較をおこなうことが可能なので、それによって信念の度合いを比べることができる。例えば、二つの命題の望ましさが等しく、(かつ、それらがともに論理的真理 (X または \bar{X}) よりも選好され)、かつ、それらの否定の望ましさも等しい場合には、それらの命題は同じ確率をもたねばならない。同様に、二つの命題の望ましさが等しく (かつ、それらがともに論理的真理 (X または \bar{X}) よりも選好され)、しかし一方の否定の方が他方の否定よりも選好される場合には、前者の確率は後者の確率よりも低い。適切な存在公理を加えると、これで、確率尺度を確立することができ、そうすればすべての命題の相対的な望ましさを確定することができる (Jeffrey, 1983, p. 113-115 / Davidson, 1980, p. 162)。

さて、デイヴィドソンの目的は意味理論を解釈することであり、意味を前提としないで信念の度合いを解釈する必要がある。それゆえ、ジェフリーの理論をそのまま使うことはできない。まず、ジェフリーのように命題が真となることを選好するという態度を解釈の基礎とすることはできない。というのも、命題とは意味なので、意味が解釈されていない段階では、行為者がどの命題を真とみなしているかがわからないからである。それゆえ、デイヴィドソンはジェフリーとは異なり、文が真となることに対する選好という行為者の態度が解釈の基礎となると考える。

そして、デイヴィドソンがジェフリーの理論を使うには、もうひとつ対処すべきことがある。信念の度合いを比較する際に、どのような選好が利用されていたのかをみればわかるように、信念の度合いの解釈のためには、文の否定や、論理的真理 (X または \bar{X}) を表す文を理解しているのでなければなら

ない。デイヴィドソンの場合、意味理論を前提としないので、これらの論理結合子と論理的真理を表す文は解釈されるべきものである。つまり、デイヴィドソンはジェフリーの理論を使うために、論理結合子（それをもとに論理的真理を表す文）を、文が真となることに対する選好をもとに解釈する方法が必要なのであり、そして彼は実際にそれを示すことに成功している（Davidson,1980,p.162-164）。以上、解釈の基礎を文が真となることに対する選好だと考え、加えてその基礎から論理結合子と論理的真理を解釈する方法を明らかにすることで、ジェフリーの理論を利用して、意味を前提とせずに信念の度合いを解釈する方法を、デイヴィドソンは明らかにした。

つまり、デイヴィドソンは、ある文が真であることよりも別の文が真であることを選好するという行為者の態度を解釈の基礎として、論理結合子、論理的真理、信念の度合い、望ましさの度合い、意味と命題的態度を解釈していく方法を示したのである。

(3) 筆者の主張

デイヴィドソンは、ある文が真となることよりも別の文が真となることを選好するという態度が、解釈の基礎となると考えた。この態度をもとに、信念の度合いが導かれ、意味が解釈可能となるという考えに問題はないように思われる。しかし、行為者がこの種の選好をもつかどうかということ、解釈者は、観察できる事実からじかに知ることができることには問題があるように思われる。つまり、選好のパターンがわかるとしたら、そこから意味を解釈することは可能だが、しかしその選好のパターンは、解釈の基礎として解釈者に利用可能なものではないのではないか。

ドレスナーは、この観点からデイヴィドソンの考えが適切ではないと考える（Dresner,2014,p.711,726）。文を真とみなすという態度を解釈の基礎としようと、文が真となることへの選好を解釈の基礎としようと、解釈者が言語的なふるまいと非言語的なふるまいを区別できると前提することになるとドレスナーは指摘する。というのも、行為者がこの文を真とみなしていると解釈者がわかるためには、解釈者は彼のふるまいが、発話という言語的なふるまいであるとわかっていなければならないからである。しかしこれは不当な前提だとドレスナーはいう。というのも、そのような言語的なふるまいは、解釈の基礎として使えるほど統一的でも一様でもなく、観察可能なものではないからである。そして、同じ問題が、文への選好を解釈の基礎とする考えにも生じるのである（Dresner,2014,p.711,726）。

ドレスナーの批判は端的に言えば、文とそうでないものを解釈者が識別で

きると前提してしまうことは不当だというものである。私は、この批判は適切なものではないと思う。確かに、(根元的解釈の状況では)ある人が音を発した時に、それが発話なのかどうかが解釈者にはわかりづらい場合もある。例えば、うめき声や、驚きの声や、雄叫びをあげるような文脈では、それが意味のない音を発することなのか、それとも発話をしているのかということがわからないかもしれない。しかし、それらの場合を除けば、ほとんどの場合には、行為者のふるまいが発話であるかどうかを解釈者にはわかると考えることは、それほど不当なことではないように思われる。そして、通常文脈においてさまざまな文を知っていくにつれて、はじめは驚きの声や、うめき声をあげる文脈でも、文や語を発しているかどうか分かるようになるだろう。また、書かれた文や語に関しては、発話かどうかという問題と比べると、わかりづらい文脈はあまりないように思われる。このように、解釈者が文とそうでないものを区別できるというデイヴィドソンの想定はそれほど不当なものではないと思われる。

しかし、解釈者が文と、そうでないものを区別することができるとしても、文が真となることへの選好を解釈の基礎とすることには、なおも問題があるように思われる。というのも、その区別ができると考えたとしても、行為者の言語的なふるまいを観察するだけでは、文が真となることへの選好をもつかどうかということは、直接に知ることはできないように思われるからである。ただし、デイヴィドソンは、どういったふるまいをもとにすれば、行為者の態度がわかるかということについて具体的なことは述べていない。行為者の態度に関するデータは、「任意のベイズ的意思決定理論の実験によるテストで通常集められるデータと同じ種類のものである」と述べる(Davidson, 1980, p261)だけである。それゆえ私は、ありうる考えをいくつか想定し、それらがうまくいかないことを示す。

言語的なふるまいをもとに選好を知ることができるかを考える前に、まずは、非言語的なふるまいから文への選好を知ることができないということを確認しておく。例えば、目の前にうどんとそばが置かれている状況で、行為者がうどんを食べたとき、そのふるまいをもとに、彼の選好を知ることができるだろうか。このとき、少なくとも行為者は、自身がそばを食べるという事態が成立することよりも、自身がうどんを食べるという事態が成立することを選好していることはわかるだろう。しかし、このとき行為者が、どの文が真となることへの選好をもっていたのかはわからない。というのも、文に関する情報は何も得られていないからである。しかし、仮にこのことがわかるとしても、この方法にはさらなる問題がある。それは、行為者が実際に真

にすることができることだけしか選択肢にあがってこないということである。例えば、行為者が雨を降らせることや、東から太陽を昇らせることはできないので、それらの事態に関わる態度については何もわからないことになる。しかし、意味の理論を解釈するという目的のためには、すべての文に対する選好関係を明らかにしなくてはならない。それゆえ、この方法は適切ではないということになる。また、非言語的なふるまいからは、論理的真理に対する態度を導くこともできないように思われる。

次に言語的なふるまいをもとに選好を知ることができるかを考える。おそらく、態度を知るために必要だと考えられているのは、提示された文の一方を選択するというふるまいとなるだろう。例えばそれは、一方の文を指さしたり、一方を手にとったり、または一方の文を読みあげるといった行為かもしれない。しかし、行為者が一方の文を指さしたという事実から、彼がその文が真となることを選好しているとみなすことはできないように思われる。というのも、その文が真となることを選好するから一方の文を指したのではなく、例えば、未来の事態について教えてあげようとして一方を指さしたのかもしれないからである。「雨が降る」と「晴れる」という二つの文を提示されたときに、晴れることの方がより望ましいと考えている場合でも、明日の天気を聞かれたと思って「雨が降る」を選択する可能性があるだろう。このように、言語的なふるまいとそうでないものを解釈者が区別できるとしても、その事実から、文に対する選好を知ることができないように思われる。

以上の考察で明らかになったことは、ある文が真となることよりも別の文が真となることを選好するという態度を、行為者がもつかどうかを、解釈者が、観察できる事実からじかに知ることはできないということである。それゆえ、文が真となることへの選好は、解釈の基礎として適切ではないということになる。すなわち、意味についてのデイヴィドソンの説明はあまり良いものではなかった。

(4) 今後の展望

デイヴィドソンの解釈理論は、基礎となる選好をもとに意味の解釈を行う方法を明らかにした。しかし、基礎としてあげられた選好は、解釈者が観察できる事実からじかに知ることはできないので、解釈の基礎と考えることはできない。では、こういった代替案が可能だろうか。ひとつは、文が真となることへの選好を解釈する方法を新たに明らかにすることで、デイヴィドソンの理論をそのまま利用することである。デイヴィドソンの理論は、文が真となることへの選好から意味の導出に関していけば問題がないように思われ

るので、この選好の解釈方法を付け加えることで、解釈理論を完成させるという方針である。もうひとつの方針はドレスナーが示している(Dresner,2014)。言語的なふるまいも、非言語的なふるまいもどちらも解釈の基礎となり、また、個人のふるまいだけでなく、言語共同体内の各ふるまいが解釈の基礎となる、という新たな考えを示している。ただし、それがどういったふるまいなのか、ということについては詳述されていない。これらの方針をより明確に示していくことが必要とされる。

(5) 参考文献

- Davidson, D. 1973, “Radical Interpretation”, In Davidson 1984a, pp. 125-139.
- Davidson, D. 1974, “Belief and the Basis of Meaning”, In Davidson 1984a, pp. 141-154.
- Davidson, D. 1980, “A Unified Theory of Thought, Meaning and Action”, in Davidson 2004, pp. 151-66.
- Davidson, D. 1984a, *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- Davidson, D. 1984b, “Expressing Evaluations”, in Davidson 2004, pp. 19-37.
- Davidson, D.2004, *Problems of Rationality*. Oxford: Oxford University Press.
- Dresner, Eli. 2014, “Decision Theory, Propositional Measurement, and Unified Interpretation”. *Mind*, 123,491, pp. 707-732.
- Jeffrey, R. 1983, *The Logic of Decision*. Chicago, IL: University of Chicago Press.
- Ramsey, F.P.1950, “Truth and Probability” in *The Foundations of Mathematics*, New York.

(九州大学)

マイノング主義論理におけるボトムアップの観点
A Bottom-up View on Meinongian Logics

小関 健太郎

Abstract

Inspired by Meinong's object theory, which is known for recognizing non-existents as genuine objects, Meinongian logics (semantics) are in a particular position between logic and ontology. Neo-Meinongian approaches of Meinongian logics are largely motivated by the need for accounting for the semantics of natural language and intentionality; while the approaches can be characterized in terms of the restrictions imposed on their Comprehension Principles, there are some problems concerning the relationship among them. In contrast to the conventional Neo-Meinongian "top-down" view, I set forth a "bottom-up" view on Meinongian logics, in which the kinds or types of objects are focused primarily so that the corresponding logics and their relations can be investigated and integrated in a piecemeal fashion.

1 研究テーマ

1.1 「論理と存在論」の問題圏におけるマイノングの対象論

アレクシウス・マイノングは、彼の対象論 (Gegenstandstheorie) を通じていわゆる非存在対象を真正の対象として認める立場を擁護したことで知られている。しばしば論点となるように、この理論はひとつの帰結として、空名辞や矛盾した記述を含む指示的表現などをそれに対応する非存在対象を指示するものとして説明することを可能にしている。影響力を持った1905年の論文「表示について」においてラッセルは、フレーゲの理論と合わせて「マイノングの理論」を——対象論を取り上げ、そうした理論はそれ自体「困難のある立場 a difficult view」であり、しかも実際にいくつかの矛盾が導かれることを指摘し、その上で記述の理論によって問題が回避されることを示した。記述の理論によるマイノング的理論の克服という考え方はクワインによっても受け継がれ、その後のメタ存在論研究のいわば「標準見解」の一部をなしてきた [1, Ch. 3]。ラッセルやクワインの議論において対象の指示 (表示) の問題が論理学の意味論的な問題に関わっているように、こうした展開は論理と存在論の関係をめぐる問題圏のうちに位置づけることができる。

マイノング自ら対象論的論理学 (gegenstandstheoretische Logik) について論じているように、マイノングにとっても対象論は論理学と密接な関係にあった [6]。ラッセルの批判以後、いくつかの契機を経てマイノングの対象論に着想を得た論理は再び注目を集めることになり、これらの諸体系はマイノング主義意味論やマイノング主義論理と呼ばれている。

本稿ではまず、こうしたマイノング主義論理の成立と展開を整理し、現代的なマイノング主義論理に伴っている固有の観点とその課題を指摘する。その上で、マイノング主義論理に対する別の見方としてボトムアップの観点を提示し、課題に対する見通しを与える。

2 研究の背景・先行研究

2.1 マイノングの対象論とラッセルによる対象論批判

マイノングの対象論のひとつの眼目は、対象概念を対象の存在や非存在によって限界づけずに「諸対象の全体」を探究することであり、このことは伝統的な意味での存在論との対比において特徴づけることができる。すなわち、存在者の学としての形而上学や存在論において「対象」はさしあたり存在者と同一視されるが、対象論において「対象」は存在者に限られず、その範囲は存在しないものにも拡大される。これによって、ペガサスのような虚構の対象や円形の四角形のような矛盾した対象も非存在対象として、対象というカテゴリーの下で実在の対象と対等に考慮されることになる。この立場は、存在論的カテゴリーなどを扱う存在論的な議論に対して、存在概念自体、また存在論や存在論的な学問についての議論や主張であるという意味でメタ存在論的な主張である。

冒頭で述べた通りこの立場は、空名辞も含め対象の指示を一様な仕方で説明する可能性を与えるが、ラッセルの指摘するこの考えの問題点は次の2点に整理できる。ひとつは矛盾の問題であり、例えば「円形の四角形」という名辞の指示対象について、それが円形であることと円形でないことが成り立つことになるが、これは無矛盾律に反する。もうひとつは存在論的身分の問題であり、例えば「存在するペガサス」が記述通りの対象であればそれは存在することになるが、われわれは任意の対象の存在を自由に操作することはできない¹。後に具体的に取り上げるように、こうした問題を回避可能なラッセルの代案がいわゆる記述の理論である。記述の理論は広く受け入れられ、1948年のクワインの論文「何があるのかについて」における存在論的コミットメントの議論でも敷衍され影響力を保つことになる。

2.2 マイノング主義論理の展開

一方でラッセルの批判の直後から、指摘された問題点に対処しうる存在論的・論理的なテクニックの発展が進み、対象論のアイデアを取り入れた論理を実現する試みもなされてきた。内的否定(述語否定)と外的否定(文否定)の区別、多値論理、自由論理などはその例であり、こうしたテクニックを基盤に、1980年前後には現代的なマイノング主義論理としていわゆる新マイノ

ング主義の諸理論が登場してきた。

なかでも自由論理は、新マイノング主義の諸理論に共通する重要な構成要素のひとつとして挙げるができる。自由論理にはいくつかの構文論と意味論が提案されているが、マイノング主義的な自由論理として、2種類の量子子の組、すなわち内的量子子 (\exists, \forall) と外的量子子 (Σ, Π) を持つものが考えられる [8, pp. 1028f.]。この体系において、例えば内的特称量子子 \exists は外的特称量子子 Σ と存在述語 $E!$ を用いて $\exists x\varphi =_{df} \Sigma x(E!x \wedge \varphi)$ と定義され、存在量子子としての内的特称量子子の「存在」へのコミットメントが明示化されることでコミットメントを含まない外的特称量子子 Σ とは区別されることになる。

2.3 新マイノング主義のモチベーション

ここで指摘しておきたいのは、新マイノング主義の多くの議論が共有する特有のモチベーションである。マイノング主義の諸理論が「自然言語の意味論と心の志向性理論の基礎」[10, p. 516] を与える試みであるとも述べられているように、新マイノング主義の取り組みの動機の基調をなしているのは自然言語や志向性の形式的な扱いを可能にするような論理の構築である。

自然言語や志向性に関する具体例として、「信じる」という志向的動詞を含む「マイノングは円形の四角形 (the round square) は円形であるということを知っていた」(1) という文を考えてみよう [4, §4]。ラッセルの記述の理論の方法を単純に適用した場合、この文は例えば「 $\exists!x (x$ は円形の四角形である かつ マイノングは x が円形であると信じていた)」のようにパラフレーズできるが、円形の四角形が存在するわけではないのでパラフレーズ後の文は偽である ($\exists!$ は一意的な存在量子子)。しかしながら元の形である (1) は実際には真なので、このパラフレーズは成功していないことになる。これに対して、マイノング的な指示理論によれば「円形の四角形」という記述は円形の四角形という非存在対象を指示するだけなので、パラフレーズなしに直接説明を与えることができる。

2.4 包括原理による対象論解釈と分類

このようなモチベーションのもとで、自然言語の意味論や志向性の問題は「表示について」以来の指示の問題と重なることになる。マイノング的な指示理論を形式化するストレートな方法は、「表示について」におけるラッセル的な対象論解釈を、「無制限の包括原理 (Unrestricted Comprehension Principle, UCP)」、つまり任意の文についてそれを満たすような対象を保証する次のような公理図式として再構成することである²。

(UCP) 任意の文 φ について、 $\Sigma x\varphi$

しかしながら、ラッセルの指摘の通り (UCP) はいくつかの問題を生み出す。そこで、なるべく寛容さは保ちつつも (UCP) に制限を加えた包括原理を与えることが求められるが、ここで (UCP) をどのように制限して寛容な包括原理を実現するかによってマイノング主義論理の分類が可能であり、代表的なものとして3つのアプローチを挙げることができる [1, pp. 110ff.]。

(二性質アプローチ) 核性質/核外性質のような2種類の性質の区別に基づいて包括原理 (以下 CP) を制限する (NCP)

(二繫辞アプローチ) 性質の例化/エンコードのような2種類の述定の区別に基づいて CP を制限する (DCCP)

(様相アプローチ) 可能/不可能世界についての様相的量化に基づいて CP を制限する (QCP)

問題となる帰結に対して、例えば二性質アプローチでは対象の「特徴」を構成するような性質 (核性質) と、存在や可能性のような様相的な性質などそれ以外の性質 (核外性質) とを区別し、穏当な制約として CP が核性質についてのみ成り立つことを認めることで存在論的身分の問題が回避される。

3 筆者の主張

3.1 新マイノング主義における「トップダウンの観点」とその課題

前節の最後では、CP の制限の方法という点でマイノング主義論理におけるいくつかのアプローチが区別されることを述べた。ここで重要なのは、前節で述べた経緯から、どのアプローチを取るにしても新マイノング主義の諸理論は——指示の理論の問題から出発して——寛容な包括原理の実現を主な関心としているという点である。このような、寛容な包括原理の実現としてマイノング主義論理を特徴づける仕方をここではトップダウンの観点と呼ぶことにする。

新マイノング主義におけるトップダウンの観点の課題として本稿で提起したいのは、この観点のもとでの各アプローチの内実と関係の問題である。まず内実の問題として、種々のマイノング主義論理ではさまざまな種類の対象が導入されるが、トップダウンの観点ではこれらの諸対象は各アプローチの道具立てに依存して定式化されている。すなわち、各アプローチの CP は、そのアプローチによる理論の枠内において非存在対象などの「マイノング主義的」対象のすべてを一挙に (アプローチによっては通常の対象も合わせて) 対象領域に導入する。一方で議論のこの段階では、その「すべて」の内部で内

実として具体的にどのような対象が導入されるのかは明らかではなく、実際この内実は非存在対象や不可能対象等々についての個別の議論によって説明されることになる。しかしながら、こうした個別の議論が CP によって導入された対象の説明をどの程度尽くしているのか、そしてその CP によって導入されていない対象があるのか、あるのであればそれはどのようなものなのか、という点は明らかではない。

このことはもう一つの、関係の問題においてより具体的に述べられる。例えば不可能や不完全対象は多くのマイノング主義論理で取り上げられる対象である。しかしながら、各々のアプローチの枠内で説明されるそれらは、同一の意味で不可能な、あるいは対象として同一な不可能対象なのだろうか？論理的に不可能な非四角形の四角形と物理的に不可能な永久機関がその不可能性に関して区別しうるように、「不可能対象」は必ずしも一義的ではなく、また不可能対象は常に同時に抽象的对象であるかどうかといった別の側面で異なる特徴づけを伴っている場合もある。また、トップダウンの観点ではどのようなアプローチが (UCP) の「適切な」制限であるかが論点であるため、各アプローチは主に問題点において比較され、両立性という点での相互関係はほとんど積極的な関心にならない³。

こうした課題を念頭に以下で概略を示したいのは、トップダウンの観点とは異なるマイノング主義論理の方法論として、ボトムアップの観点を導入することである。

3.2 ボトムアップの観点の導入

ボトムアップの観点の基本的なアイディアは、対象領域への新しいタイプの対象 (または対象のタイプ) の導入と、それによる対象領域の構成あるいは拡張としてマイノング主義論理を特徴づけることである。

こうした方法論は、関連研究においても部分的に見い出すことができる。第一に、マイノングの対象論それ自体が最も重要な源泉として挙げられるだろう。マイノングは対象論をいくつかの仕方で特徴づけているが、存在論の拡張としての対象論の特徴づけは、まず実在的对象の領域からイデア的对象を含めた領域への拡大から出発し、さらに進んで非存在対象やその下位区分へと拡大される形で論じられている [5]。新マイノング主義のうちでボトムアップの観点に特に接近しているのは [12] で再検討されているシルヴァンのアイテム理論 (item theory) であり、CP に代わる「公準」(postulates) などによる対象の導入と、それに応じた対象領域の区分けという方向性のもとでトップダウンの観点の延長線上にあるものの、ボトムアップの観点同様に特定の CP によらず対象の導入を認めることが主張されている。

「対象の全体」という対象論的な見取り図を共有しつつも、全体そのものをカバーする (UCP) をどのように制限するべきかが論点となるトップダウンの観点に対して、ボトムアップの観点では対象の様々なタイプに応じてその特徴づけを与え、それらに関係づけていくことで構成的に全体をカバーすることが目指される。重要な点として、ボトムアップの観点はトップダウンの観点と排反なものではない。そこで以下ではボトムアップの観点の具体的な適用として、先に挙げたトップダウンの観点における内実と関係の問題を取り上げる。

3.3 マイノング主義論理の構成要素

まず内実の問題に関して、対象のタイプを関心とするボトムアップの観点において注目されるのは、既存のマイノング主義論理においてどのような対象のタイプが扱われているのかという点である。実際にさまざまな対象のタイプが先行研究から収集されるだろう——非存在対象や不可能対象、志向的对象、抽象的对象、また不完全対象などのギャップ対象はその例である。

このうちギャップ対象 (gappy objects) を取り上げよう。ギャップ対象は、その対象についての文 A に関して A も A の否定も成り立っていないような対象である。例えばシャーロック・ホームズという人物は、「シャーロック・ホームズは背中にほくろがある」もその否定も成り立っていないという点でギャップ対象として考えることができる [9, p. 183f.]。二性質アプローチや二繫辞アプローチのもとでは、ギャップ対象は基本的に外的否定と内的否定や述定の区別を利用して矛盾対象と実質的に同じ仕方として導入されている (cf. [9, p. 106])。これに対して、CP の制限方法に依存せずギャップ対象そのものに注目した場合、他の仕方でもギャップ対象を導入することもできる。例えば、[11] はある種のギャップ対象であるマイノングの不完全対象に言及しつつ、ギャップ対象の論理について異なる形式化を与えている。また、否定を区別するのではなく加法的選言と乗法的選言のように 2 種類の選言を区別することや、超付値 (supervaluation) のような方法も候補となる。場合によってはこれらはギャップ対象の異なる導入の仕方というより、ギャップ対象というタイプの細分化でもありうる。

3.4 非対立的なマイノング主義論理

この延長線上において、マイノング主義論理の各アプローチがどのような関係にあるのかという関係の問題にも見通しが与えられる。トップダウンの観点の問題として、この観点ではあくまでどのようなアプローチが (UCP) の「適切な」制限であるかが論点であるために各アプローチの相互関係が積極的な関心にならないことを論じたが、ボトムアップの観点では各アプローチが

導入する対象に注目してアプローチ間の関係を調べることができる。

例えば、(NCP) と (DCCP) を比較することを考えてみよう。比較のひとつの方法は、あるタイプの対象に注目し、それが各アプローチの CP によって導入される対象のうちに含まれるかどうかを比較することである。実際、(NCP) と (DCCP) によって導入される対象のタイプには、ギャップ対象のように実質的に重複するものもあれば、(DCCP) を採用する [13] における「抽象的对象」のように一方に固有のものも見られる。こうした比較を通じて、それぞれの CP のカバー範囲の重複や差を含め、CP の位置付けや相互関係を調べることができる。関連して [12] は、(QCP) と同様の現実世界とその他の世界の (意味論的な) 区別の下で、その他の世界に存在する対象を現実世界へ非存在対象として取り込む方法として (NCP) のような CP を用いることができることを指摘している。

4 今後の展望

新しい観点を導入してマイノング主義論理を再検討することは、論理と存在論の両側面できざまな対象についてより柔軟な分析を可能にする。ここでは「より柔軟な分析」の内実に関して二つの展望を述べておきたい。

ひとつは、マイノング主義論理というプロジェクトの内部における諸対象の分析であり、この点については、種々の理論によって定められる対象領域の間の関係や、複数の CP の間のメタ的な関係など、新しい観点によって可能になるさまざまな問題設定が [12] において近い立場から素描されている。

もうひとつは、マイノング主義論理をより広い文脈に位置づけることによるものである。マイノング主義論理はこれまで必ずしも十分に他の論理との関係の面で考察されてきていないが、ボトムアップの観点によってマイノング主義論理を部分的に取り上げることが可能になることで、すでに自由論理や矛盾許容論理といった側面に関してある程度なされているように、より広い哲学および数理論理学の文脈における考察を通じて、マイノング主義論理内外のいわば相互交流が実現されることが期待される。

注

¹ 現実にはペガサスは存在しないことが成り立っているので、ラッセルはこの問題も矛盾の問題として扱っているが、この問題は無矛盾律に依存しない。他にも関係に関する問題が挙げられる [2]。

²[1, p. 108] に従う。シルヴァンらはほぼ同じ意味で特徴づけ原理 (Characterization Principle) という語を用いている。

³[3] は二性質アプローチと二繫辞アプローチの関係についての論争を反映しており、論争に関わる論文と合わせて、批判的ではあるが相互関係についての考察を含んでいる。

文献

- [1] Berto, F. & Plebani, M., 2015. *Ontology and Metaontology*. Bloomsbury.
- [2] Griffin, N., 2009. “Rethinking Item Theory”. *Russell vs. Meinong: The Legacy of “On Denoting”*. Routledge, 204–232.
- [3] Jacquette, D., 1997. “Reflections on Mally’s Heresy”. *Axiomathes*, 8(1), 163–180.
- [4] Marek, J., 2013. “Alexius Meinong”. In: Zalta, E. N. (Ed.). *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), URL=<<https://plato.stanford.edu/archives/fall12013/entries/meinong/>>.
- [5] Meinong, A., 1904. “Über Gegenstandstheorie”. In: *GA II*, 481–530.
- [6] Meinong, A., 1910. “Erstes Kolleg über gegenstandstheoretische Logik”. In: *GA Ergänzungsband*, 209–236.
- [7] Meinong, A., 1969–1978. *Gesamtausgabe*. Akademische Druck- und Verlagsanstalt. [=GA]
- [8] Nolt, J., 2005. “Free Logics”. In: Jacquette, D. (Ed.) *Philosophy of Logic*, Elsevier, 1023–1060.
- [9] Parsons, T., 1980. *Nonexistent Objects*. Yale University Press.
- [10] Rapaport, W. J., 1991. “Meinong, Alexius I: Meinongian Semantics”. In: Burkhardt, H. & Smith, B. (Eds.). *Handbook of Metaphysics and Ontology, Volume 2*, Philosophia, 516–519.
- [11] Santambrogio, M., 1990. “Meinongian Theories of Generality”, *Noûs*, 67, 21–36.
- [12] Sylvan, R., 1995. “Re-Exploring Item Theory. Object Theory Liberalized, Pruralized and Simplified but Comprehensivized”, *Grazer Philosophische Studien*, 50, 47–85.
- [13] Zalta, E. N., 1983. *Abstract Objects*. D. Reidel.

(慶應義塾大学)