

Title	解釈と行為の説明における寛容の原則
Sub Title	The principle of charity in interpretation and the explanation of action
Author	宮館, 恵(Miyadate, Satoshi)
Publisher	三田哲學會
Publication year	1988
Jtitle	哲學 No.86 (1988. 6) ,p.1- 26
JaLC DOI	
Abstract	Davidson and others have recommended principles of charity. In this article I am defending their claim that interpretation and the explanation of action are impossible without a rather strong principles of charity. The principle of charity in the domain of interpretation is that when we interpret what others say we must assume their beliefs are usually true. The principle of charity in explaining action says that the speaker must generally be rational both in the sense that his beliefs must be logical and in the sense that his action must be reasonable given his desires and beliefs. These principles should be considered as constitutive for interpretation and the explanation of action by using such concepts as desires, beliefs and other attitudes.
Notes	
Genre	Journal Article
URL	<a href="https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00150430-00000086-0001">https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00150430-00000086-0001</a>

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the Keio Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

# 解釈と行為の説明における 寛容の原則

— 宮 館

恵\*

## The Principle of Charity in Interpretation and the Explanation of Action

*Satoshi Miyadate*

Davidson and others have recommended principles of charity. In this article I am defending their claim that interpretation and the explanation of action are impossible without a rather strong principles of charity.

The principle of charity in the domain of interpretation is that when we interpret what others say we must assume their beliefs are usually true. The principle of charity in explaining action says that the speaker must generally be rational both in the sense that his beliefs must be logical and in the sense that his action must be reasonable given his desires and beliefs.

These principles should be considered as constitutive for interpretation and the explanation of action by using such concepts as desires, beliefs and other attitudes.

人を理解する一つの方法は、その人の行動を説明することである。その人が何を考えてなぜその行動をしたのかという説明を与えることは、確か

---

\* 慶應義塾大学大学院文学研究科博士課程（哲学）

にその人を理解することに寄与するであろう。ところが、人の行動を理解するには、その重要な一部として、その人の発話行動を理解しなければならない。人が何か言葉を発したとき、それが我々の知っている言語であれば、我々はその意味を理解する。知らない言語であれば、それを知りたいと思う。おそらく、言葉の理解にかかわっている事柄は、いずれの場合も同じであろう。どのような知識があれば、我々は人の発話を理解できるであろうか。

ディヴィッドソンは、真理条件的な意味理論によって、言語理解の理論を提示しようとしている。それを詳細に検討してみると、そこでは「寛容の原則」という独特の興味深い原理が用いられていることがわかる。これは単一の原理ではなく、用いられている場面によって大きく分けて二つの異なった内容を持つ。一つは、人の信念をできるだけ真とせよという「真理に関する寛容の原則」、もう一つは、人をできるだけ合理的とせよという「合理性に関する寛容の原則」である。

言葉の意味にかかわることは、最終的には行動の説明そして人の理解という脈絡に位置づけられなければならない。本稿は、その観点から、この原理がどのような性格のものであるかを考察しようというものである。Ⅰでは、ディヴィッドソンが提案している解釈理論を概観し、そこに仮定されているいくつかの重要な条件に注目する。Ⅱでは、それらのうち特に意味解釈にかかわるものを考察し、真理に関する寛容の原則がその企画にとって構成的原理であることを確認する。Ⅲでは、日常的に行なわれる目的論的な行為の説明が、二つの意味の合理性に関する寛容の原則を用いており、そこでもやはりそれは構成的原理として働いていることを見る。Ⅳでは、解釈や行為の説明は、信念、欲求、意図等の概念を用いる限り、寛容の原則を使用せざるをえず、その使用によって一般に特徴づけられることを示す。ディヴィッドソンの意味論については、もっぱら寛容の原則の使用という側面から考察し、真理条件的意味論の妥当性については論じない。

なお、本稿は、日本科学哲学会第20回大会（昭和62年11月7,8日、上智大学）における同名の研究発表に加筆したものである。

# I

一般に、自然言語に属する文は無限である。いわゆる単語などの基本的な語彙は有限個であると考えても、それらから作られる文は無限に可能である。<sup>(1)</sup> もちろん、現実人間どうしの会話で使用される文としては、その長さや構造上の複雑さの点で確かに制限がある。しかし、解釈者は、対象言語に属する無限個の文を原理的には解釈することができなければならないから、個々の発話を一つ一つ集めていってそれぞれに解釈を付与し、その総体を解釈理論とするわけにはいかない。そこで我々は、個々の文の解釈を引き出すことができるような有限の理論を述べなければならない。

ディヴィドソンの考えている第一階の言語  $L$  に対する真理理論 ( $T$ -理論) は、 $L$  の単称名辞の指示対象と一般名辞の外延とを特定する有限個の公理、および、 $L$  の文に対する真理の回帰的定義から成る。 $T$ -理論は、 $L$  の各文  $s$  について、

(1)  $s$  は ( $L$  において) 真である iff  $p$ .

の形式の  $L$  に対する真理述語を含むメタ言語 (ここでは日本語) の文 ( $T$ -文) を含意しなければならない。ここで  $s$  は  $L$  の文の正規的な記述 (名前)、 $p$  はメタ言語の文である。内容的に適切な  $T$ -理論は、 $L$  の各文  $s$  について、 $T$ -文の双条件の右辺の文  $p$  が左辺に記述された文  $s$  の正しい翻訳となっていなければならない。これは規約  $T$  (Convention  $T$ ) と呼ばれる。すなわち、規約  $T$  は、両辺の文が (直観的な意味で) 同義であることを要求している。

たとえば、対象言語を英語、真理理論を述べるメタ言語を日本語 (アルファベット等を含む) とすると、

(2) 'Snow is white' は真である iff 雪は白<sup>(2)</sup>い。

この  $T$ -文は、「'Snow is white' という英語の文は、雪が白いとき、そしてそのときに限り真である」ということを日本語で述べている。すなわち、英語の文 'Snow is white' の真理条件を日本語で与えている。右辺の  $p$  の位置には左辺の  $s$  の翻訳文が入っているのだから、その正しさは、英語から日本語への翻訳理論の正しさに依存している。

また、対象言語を日本語、メタ言語を同じく日本語とすると、次のような  $T$ -文が得られる。

(3) 「雪は白い」は真である iff 雪は白<sup>(3)</sup>い。

これは、「『雪』『は』『白』『い』という日本語の連鎖は、雪が白いとき(一般に雪とは白いものであるとき)、そしてそのときに限り真である」ということを述べている。この場合も、 $p$  は  $s$  の翻訳になっていなければならないのだから、この  $T$ -文が規約  $T$  を満たして適切な真理理論からの定理となっているかどうかは、日本語から日本語への翻訳理論に依存していることになる。<sup>(3)</sup>

このように、真理理論から導出されるそれぞれの  $T$ -文は、タルスキにおいては、iff の右辺が左辺の  $s$  の翻訳文であるとされているから、翻訳が正しい限り、 $T$ -文全体は真であることが保証されている。つまり、翻訳を仮定して真理が規定できた。しかし、我々が行なおうとしているのは根底的解釈なのであるから、前もって正しい翻訳が得られていると仮定することはできない。そこで、ディヴィッドソンの提案は、逆に真理を基本と考<sup>(4)</sup>えて、翻訳や解釈に関する説明をそこから引き出そうということである。つまり、翻訳関係を仮定せずに、なんらかの経験的な根拠によって  $T$ -文が真であることがわかれば、左辺の対象言語の文  $s$  の解釈として右辺の  $p$  を使用できるのではないか、<sup>(5)</sup>ということである。そこで、タルスキの同義性の要求(規約  $T$ )は、翻訳関係すなわち意味の概念に頼らない形で、次

のように言い替えられる。

規約  $T'$  : 受容可能な真理理論は、対象言語のそれぞれの文  $s$  について、次の形式の定理を含意しなければならない。

$s$  は真である iff  $p$ .

ここで ' $s$ ' には  $s$  の正規的な記述が入り、' $p$ ' には  $s$  が真であるときそしてそのときに限り真であるような (メタ言語の) 文が入る。<sup>(6)</sup>

このように言い替えられた規約  $T'$  は、元の規約  $T$  と異なり右辺と左辺の同義性を要求しているのではない。そうではなくて、適切な  $T$ -理論から導出されるすべての  $T$ -文が真であることを要求しているのである。それが満たされるならば、右辺の  $p$  を左辺の  $s$  の解釈として使えるかもしれないからである。したがって、規約  $T'$  は、真理理論の内容的な適切さに関する条件を、 $T$ -文を通していわば間接的に述べていることになる。これは、導出される  $T$ -文こそが、経験的なチェックを受けるからである。そこで、ディヴィドソンの主張をまとめると、内容的および全体論的制約に従う真理理論は、すなわち自然言語の意味論である、ということになる。これは Davidson (1967) で提示された彼の最初の立場である。

その後、ディヴィドソンはその立場に修正を施し、言語を理解するのに十分な知識は何かという問を立て、その答えとして、「理論を知っている人なら、それが適用される発話を解釈できる」ような理論、という言い方を<sup>(7)</sup> する。すなわち、ある言語を理解するのに十分な、その言語の話者の持っている知識に言及して  $T$ -理論の利用法が述べられている。つまり、内容かつ全体論的に適切な  $T$ -理論は、それ自体では意味論ではないが、人が、 $T$ -理論を知り、かつ、その  $T$ -理論が形式的小および経験的規準を満たすことを知るならば、その人は、導出される  $T$ -文をその言語の発話を<sup>(8)</sup> 解釈するために使用することができるとされる。これは、最初の定式化にいわば認識論的な転回を与えて改訂したものである。

さて、我々の最終的な目標は、人間の言語的および非言語的行動を説明することである。ある人が何を考え、どういう理由である行動をしたのか(するのか)ということを説明するために、我々はその人の話す言葉を理解したいと思う。つまり、その人の話す言葉を我々の知っている言葉で解釈したいと思っている。我々の求めている解釈理論は、この意味で、人間の行動を説明する大きな理論の一部として機能すべきものである。

話者は、ある事柄についてなんらかの信念を有し、その信念を文に託して表現する。たとえば、今外で雨が降っていると話者は考えているとする。それをだれかに伝えるとき、彼は、彼の言語で雨が降っているということを意味するなんらかの文を発話する。聞き手がその言語を理解できる人であれば、その発話によって言われた事柄を理解し、彼の発話行動を解釈する。すなわち、今外で雨が降っているということを彼は彼の言語で言っているのだということを知る。話者は、ある事柄を他人に伝えたいという理由で、あるいはある事柄を他人に隠したいという理由で、あるいはその他なんであれなんらかの目的で、彼の言語で何かを意味する言葉を発する。我々は、その発話行為を何かを言う行為として解釈したい。そこで問題は、話者の信念についても話者の言語の意味についてもあらかじめ知らないときに、どのようにしたら話者の発話行動を解釈できるだろうか、ということである。これは、クワインの根底的翻訳 (radical translation) にならって、根底的解釈 (radical interpretation) と呼ばれる<sup>(9)</sup>。

話者の発話を解釈するとは、その発話行為がどんな事を言っている行為であるかを述べることである。したがって、それは、話者が自己の言語の文を断言することによって表明しているところの信念を特定することである。彼の 'It's raining' という発話行為を、今雨が降っているという信念を表明する行為として観察し、その発話された文と表明された信念とを体系的に関連づけることで、我々はそれらの発話の意味を推測することができる。そこで、話者の発話の意味を知るためには、話者の信念を知らなけ

ればならない。ところが、話者の非言語的な行動を見て、彼が信じていることを明らかにするのは、一般に不可能である。たしかに、かなり大まかな信念ならば、話者の外的行動から推測できるかもしれない。たとえば、外で雨が降っていると思っているかどうかぐらいは、雨音が聞こえる位置にいるとか、あるいは雨の降る様子が実際に見える位置にいるとかいう状況の下で、いつもと違ってなかなか外へ出ようとしないとか、ちょっと出てはすぐぬれない所へ戻るとかいう行動から、ひょっとすると判断できるかもしれない。

しかし、もっと詳しく、たとえば、話者が今日の雨は午前中で止むだろうと思っているのか、それとも夕方まで降ると思っているのか、また、この雨は雲から降って来ると思っているのかそれとも雨の神が降らすと思っているのか、農業にとって恵みの雨と思っているのかいないのか、という具合にその信念内容がある程度細かく特定していこうとすれば、たちまち話者自身の言語的行動、すなわち言葉による信念の説明に頼らざるをえなくなる。ほとんどすべての信念は、言葉によってはじめて分節化されるのである。発話の意味を知るためには話者の信念を知らなければならないが、その信念を知るには発話の意味を知ることができなければならない。人の意図や信念を詳細に理解することは、その人の発話を理解することと独立ではありえず、どちらか一方についての理論を他方の前提とすることはできない。<sup>(10)</sup>

意味と信念のこの循環を断ち切るために、ディヴィッドソンは、話者が自己の文を真と考える、という態度を基本に置くことを提案する。<sup>(11)</sup>これはもちろん信念の一種ではあるが、特に意味の問題とは切り離して特定することのできる信念である、とディヴィッドソンは主張する。話者は、自己の信念の表明として、ある状況の下で‘It’s raining’と発話する。そのとき、その話者はその状況の下で‘It’s raining’という自己の文を真と考えている、と我々は通常推測できよう。そこでさらに、その信念が成立している



又は成立していない 外的状況を 記述することができれば、我々は ‘It’s raining’ の文の意味を知らずに、また、その未知の話題についての話者の信念も、その発話に 関与する話者の詳細な意図も知らずに、話者にとっての ‘It’s raining’ の真理条件を与えることができることになる。あるいは、 $T$ -理論の検証という 観点から言うなら、 $T$ -理論から 定理として導出される具体的な  $T$ -文が正しいかどうか、すなわち規約  $T'$  が満たされているかどうかを経験的な証拠に基づいて判定できる ことになりそうに思われる。

しかし、注意深く検討すべきはまさにこの点である。 $T$ -文が与えられているものは、対象言語  $L$  の文  $s$  の真理条件なのであって、 $L$  の特定の話者たちが  $s$  を真と考えるときの外的状況（真理条件）ではないはずである。また、右辺に現われる真理条件  $p$  も、解釈者が文 ‘ $p$ ’ を真と考える状況ではなく、端的に  $p$  である状況のはずである。つまり、

(1)  $s$  は ( $L$  において) 真である iff  $p$ .

であったはずのものが、

(4) ( $L$  の) 話者は  $s$  を状況  $C$  の下で真と考える iff 解釈者は  $C$  の下で ‘ $p$ ’ を真と考える

という形になっているのである。

これをもう少し詳しく見てみよう。ある英語の話者は、ある状況  $C$  の下である信念を抱いて、‘It’s raining’ と発話する。彼は  $C$  の下で ‘It’s raining’ を真と考えている（と解釈者は推測する）。さて解釈者は、 $C$  の下で「雨が降っている」という日本語の文を真と考える。すなわち、解釈者は  $C$  の下で、雨が降っているという信念を抱く。意味を知るためには話者の信念を知らなければならないから、解釈者は  $C$  の下での自己の信念を話者に帰属させる。すなわち、 $C$  の下で話者は雨が降っているという

信念を抱いている、と解釈者は想定する。そこで、話者の‘It’s raining’  
という発話は、雨が降っているというその帰属させられた信念を表明した  
ものだ、と解釈者は推測する。したがって、 $T$ -文が<sup>(12)</sup>

- (5) 話者は  $C$  の下で ‘It’s raining’ を真と考える iff 解釈者は  
 $C$  の下で「雨が降っている」を真と考える

となっていれば、信念の共有を条件として、この (5) が真であることを経験的に確認できることになる。

ところが、これらの一連の手続きを進行させるには、さらにいくつかの重要な条件を補う必要がある。まず第一に、信念と発話の関係であるが、発話から信念が推測できるためには、我々は、話者の発話のほとんどは彼が真と考える文の発話である、と仮定しなければならない。そうでなければ、発話された文から信念を推測することはできない。これは「信念表明の正確さ」の条件と言えよう。一般に、話者が  $p$  と言うならば話者は  $p$  と信じている、と仮定しなければならない。

第二に、我々は、話者が ‘It’s raining’ を真と考えているときの外的状況  $C$  (話者にとっての真理条件) を記述できなくてはならない。したがって、当面扱われる文は、外的状況の中の発見可能な変化に応じてその真理値が変動するような文である。クワインの用語で言えば、「観察文」または一般に「場面文」である<sup>(13)</sup>。そうすると、その文の真理値は、発話された時刻、場所、発話者に相対的に決まるのだから、真理理論はこれらのものに相対化されなければならない。この相対化は、自然言語の指示的要素を扱うことができるようにするためでもあると同時に、そもそも解釈理論を得る経験的な手がかりとして当然仮定されなければならない。

第三に、解釈者が自分の信念を話者に帰属させるためには、一般に、話者がある文を真と考えるときと、その文が真であるときとが一致すると仮定しなければならない。強い言い方をするなら、すべての話者は、文が真

であるときそしてそのときに限りその文を真と考える、という仮定である。<sup>(14)</sup> 文が真であるかどうかは、結局は解釈者の信念の下での話であるから、さらに正確に定式化すれば、話者は、解釈者が自己の信念に照らして文が真であると考えるとき、そしてそのときに限りその文を真と考える、ということになる。第一の「信念表明の正確さ」の仮定により、発話された文は話者の信念を表わすから、結局、話者の信念と解釈者の信念は一致することになる。これが保証されるならば、話者によって真と考えられている文がどんな信念を表わすかを我々は知ることができるので、我々は、彼の発話をそのかくかくの信念を表明する行為として解釈することができる、つまり彼の文を解釈することができることになる。

この仮定は、信念を媒介として発話や文の解釈を行なおうとするディヴィッドソン流の意味理論にとって最も重要な仮定であって、信念を話者に帰属させるときの基本的原則である。

(理論と、理論構築者の事実に対する観点に従って) 文が真であるとき、できるだけ話者にそれらの文を真と考えさせるように、真理条件<sup>(15)</sup>を選ぶ。

解釈者と被解釈者の信念の一致を要請するこの仮定は、(真理に関する) 寛容の原則 (Principle of Charity) と呼ばれる。解釈者は被解釈者に自己の信念を投影して一定に保ち、それによって意味と信念の相互依存性の問題を解こうとするものである。<sup>(16)</sup>

上で述べた第二点、すなわち真理理論を時刻、場所、話者に相対化するという点、および第三点、すなわち彼我の信念を一致させるという寛容の原則、の二つのことから、前述の (4) は、*T*-文の型ではなく、むしろ、*T*-文 (1') の *s* と *p* に関する条件と考えるべきである。

(1') *s* は (言語 *L* において) 話者 *x* にとって時刻 *t*, 場所 *l* において真である iff *p*.

その具体例は、

- (6) ‘It’s raining’ は英語において話者  $x$  にとって時刻  $t$ , 場所  $l$  において真である iff 時刻  $t$ , 場所  $l$  において雨が降っている。

となる。このとき、‘It’s raining’ という英語の文と「雨が降っている」という日本語の文は、(4) を満たすように選定されており、さらにもし規約  $T'$  を満たすならば、この  $T$ -文 (6) (および他のすべての英語の文についても同様に) を含意する  $T$ -理論は内容的に正しいと経験的に判定される。

そこで、第四点として指摘すべき条件が明らかとなる。それは、(4) を満たすように選定された文  $s, p$  が、規約  $T'$  に述べられた  $s, p$  の条件を満たすための仮定である。すなわち、話者および解釈者が真と信じているということから、その文が真であることへの移行を可能にするような仮定である。一般に、 $x$  が  $s$  を真と信じているならば、 $s$  は真である。 $T$ -理論を話者に相対化したのだから、この第四の条件は不要であると思われるかもしれない。しかし、ディヴィッドソンの話者への相対化は、主として指示的要素を扱うためのものであるから、その相対化とは別に、信念帰属の要件としてやはりこの第四の条件は必要である。<sup>(17)</sup>

次節では、他人の発話の解釈理論を構築するためのこれらの条件について、その性格、すなわち、経験的な指標としての条件であるのか、あるいはまた、なにかこの企画に特有の非経験的な原理であるのか、という問題について考察する。

## II

寛容の原則 (Principle of Charity) という言葉は、Wilson (1959) “Substances without Substrata” に由来する。この論文で彼は、名前の

記述理論について論じている。チャールズという人物が、「シーザー」という名前を含む次の五つの断言を行なったとしよう。<sup>(18)</sup>

- (i) シーザーはゴールを征服した。
- (ii) シーザーはルビコン川を渡った。
- (iii) シーザーは3月15日に暗殺された。
- (iv) シーザーは独立奪格の使用に凝っていた。
- (v) シーザーはブーディカと結婚した。

そこで、「シーザー」という名前の指示対象としてどの個体を選ぶかという問題に関して、ウィルソンは次のように述べている。

したがって我々は、寛容の原則とでも呼べるものに基づいて事を行なう。我々は、指示対象として、できるだけ多くのチャールズの言明を真とするような個体を選定する。この場合、それはジュリアス・シーザーという個体である。我々は、その指示対象は、「シーザー」という語を含む断言された五つの文を、他のいかなる個体よりも多く満足するところの個体である、と言ってよからう。<sup>(19)</sup>

すなわち、チャールズの五つの言明をできるだけ真と考えて、そうなるように名前の指示対象を選ぼう、ということである。

寛容の原則という言葉は、それ以来さまざまな場面で用いられて来ているが、その代表例はIで述べたように、発話の解釈の場面で用いているデヴィッドソンである。彼は、後述するように、文の解釈に限らず行為の説明の場面においても寛容の原則を使用しており、その場合の「寛容の原則」は、解釈での真理に関するものとはまた違った意味を持っている。また、翻訳の問題に関してクワイン<sup>(20)</sup>、心理学の方法論に関してデネット<sup>(21)</sup>、ソーバー<sup>(22)</sup>、コーエン<sup>(23)</sup>、スティッチ<sup>(24)</sup>、あるいは、経済学や決定理論においても<sup>(25)</sup>、同様の問題が議論されている。

これらは、大きく分けて次の二つの種類に分類することができる。一つは真理に関するもので、他人の言葉の意味を解釈するとき、被解釈者の信念をたいていの場合真であると仮定する「真理に関する寛容の原則」である。我々は、他人の信念がたいてい真となるようにその人の発話の意味を定めなければならない、ということである。もう一つは、合理性に関するもので、行為を説明するとき、行為者はたいていの場合合理的であると仮定する「合理性に関する寛容の原則」である。我々は、行為者がたいていの場合合理的となるように、その人に信念や欲求を帰属させなければならない、ということである。

これら二つの種類の寛容の原則は、全く独立して使用されるわけではなく、互いに交差して使用されているのであるが、発話の意味解釈の場面では主として「真理に関する寛容の原則」が重要であり、行為の説明の場面では「合理性に関する寛容の原則」が強調される。これらの原則の地位や強さの程度は、論者によってさまざまである。経験的仮説であるという者<sup>(26)</sup>、しかもその大部分は偽であるとする者<sup>(27)</sup>、また進化論の立場から正しさを主張する者<sup>(28)</sup>、人類学的考察から疑義を唱える者<sup>(29)</sup>、など活発に議論されている。以下では、これらの寛容の原則は、解釈や行為の説明にとって、信念・欲求・意図等の概念を使用する際のルールを規定する構成的原理であることを論じる。

I で見たように、解釈の場面では、二つの大きな仮定が使用されていた。一つは、話者の発話はたいていの場合彼の信念の表明であって、信念を正しく反映するものだという「信念表明の正確さ」の仮定、もう一つは、話者の信念はたいていの場合真であるという「真理に関する寛容の原則」の仮定であった。これらは一見すると明らかに誤りであって、反例の枚挙に困らないように思われる。真理値を持ちうるタイプの文に限っても、うそは、真実と違ふと知りつつあえて言うことだし、間違って真実だと信じて発話することもよくあることだ。同じ言語の話者でも、政治や宗

教といったある種の分野においては、人によって意見が異なるのがむしろ普通である。科学的な研究の分野でさえ、意見の衝突が見られるのは決して珍しくない。だとすれば、発話は必ずしも信念を正確に反映しないし、まして人の信念がたいてい真だとはとてもありそうには思われない。したがって、人間の発話を解釈しようというときには、むしろ、「信念表明の不正確さ」や「信念の可謬性」をこそ念頭に置いて、探究を進めるべきではないかとさえ思われるかもしれない。

しかし、「信念表明の正確さ」や「寛容の原則」は、探究を進める際の発見的指針として提案されているのではないのである。その二つを一応の指針として立てれば、そうしないよりも発話の解釈がうまくいくだろうと言われているのではない。そうではなくて、信念を帰属させて解釈を行なおうという試みそのものを、そもそも成り立たせるような「構成的原理」(constitutive principle)としてそれらは主張されているのである。

もし、……すべての話者は、文が真のときそしてそのときに限りその文を真と考える、と仮定すれば、……すなわち、少なくとも表現される限りすべての信念は正しい、と仮定すれば、正しい（解釈）理論を保証することが可能となるであろう。

しかし、もちろん、話者が偽なる信念を決して持たないとは仮定できない。間違いがあるからこそ、信念には信念としての効用があるのだ。しかしながら、我々は、大部分の信念は正しいというのは所与のことだと考えることができる。その理由は、信念は、諸信念のパタン(30)におけるその位置によって同定されるからである。……

すなわち、信念という概念を使用して解釈を行なうには、その基本的な前提として、大部分の信念の一致を「ア・プリオリに」(31)認めなければ話が始まらないというのである。被解釈者や解釈者は、いずれも、誤りすなわち偽なる信念を一切持たないわけでは決してない。ただ、不一致や誤りを解

釈するには幅広い一致が不可欠だ、幅広い一致を背景としてはじめて不一致や誤りはその意味を持つようになる、ということである。<sup>(32)</sup>

そこでのポイントは、他人の言葉（そして後述のように行為も）の理解は、ある種の基本的な一致を背景としてのみ可能だ、ということであり、したがってその理解に寄与すべきものである解釈理論は、理解の基盤となる一致を提供する形で他人の言葉の意味を解釈しなければならない。ここで得られる一致は、とりもなおさず我々が持っている信念体系との一致であるので、結局、我々の目から見て真であると思われる大量の信念を相手の中に見出す、つまり話者に帰属させることによって、この一致は得られることになる。

もし、我々自身の基準から見て大体整合的で真なる信念体系を示すものとして、ある生物の発話や行動を解釈する方法が発見できなければ、我々はその生物を、合理的として、信念を持つとして、およそ何かを言うとして見なす根拠がなくなってしまう。<sup>(33)</sup>

このようにして、信念と意味を帰属させることによってその話者の言語を理解しようという解釈の企画にとって、「話者の信念の真理性に関する寛容の原則」は、解釈を上手に行なうための方針というよりは、この企画そのものを成り立たせるところの構成的原理であると考えられる。

この議論の基本的な根拠は、信念と意味は人の心の中で確定して分離しているわけではないという信念・意味の全体論と、理解には大量の一致の基盤が必要である、というこの二つである。

全体論との関連で、クワインの翻訳の不確定性のテーゼを考慮して付け加えるならば、解釈理論もまた、信念と意味の分離のさせ方に応じて同様に不確定である。信念と意味の全体論は、常に存在論的な主張であって、なにか証拠不足のために分離が確定しないということではない。そうではなくて、信念と意味は上で見たように解釈において互いに補完的な働きを



しているため、ある信念を話者に帰属させるということがすなわちある仕方  
方でその人の文を解釈することであり、他の信念をその人に帰属させる  
ということがすなわち他の仕方方でその文を解釈することになるのである。  
「真理に関する寛容の原則」は、解釈者と被解釈者の信念のかなりの一致  
を保証することによって、この不確定性の幅を狭めるのではないかと考え  
られるかもしれないが、そうではない。<sup>(34)</sup> この原理は、解釈理論の成立その  
ものについての構成的原理なのであって、でき上がった解釈理論を比較し  
たり選択したりするときの何か規準を提供するといったものではないから  
である。

「全体論」という言葉で意味されていることが、信念・意味の全体論<sup>(35)</sup>  
であれ、文の経験的意味はその文を含む理論全体の中で決まるという全体論  
であれ、いずれにせよ全体論からは必ず不確定性が帰結する。全体論の立  
場では、文への意味の割り振り方をどこかで変更しても、その変更は他の  
箇所では何らかの調整を施すことによって、全体としては経験的証拠に反し  
ない形で相殺されうるからである。全体論のポイントは、解釈や行為の説  
明で用いられる信念・意味・欲求等の概念の道具立てが独立した経験的  
内容を持たず、全体としてのみ経験的価値が与えられているというところ  
にある。

### III

以上では、ディヴィッドソンらによって「寛容の原則」と言われているも  
ののうち、特に真理に関するものについて述べて来た。これは、主として  
他人の言葉の解釈にかかわる場面で主張されており、構成的原理という性  
格を持つ。ところが、同じ「寛容の原則」とはいっても、行為の説明の場  
面で用いられるときには、また異なった内容が込められている。それは、  
他人をできるだけ合理的であるとして説明せよという「合理性に関する寛  
容」の原則である。

他人の行為を理解一つの方法は、その行為について説明を与えること、すなわち、なぜその行為がなされたかという理由を述べることである。我々の常識的な心理理論とでも言うべき folk psychology (常識心理学) によれば、この説明は、たとえばおよそ次のように行なわれる。

もし、

ある人が新鮮な空気を部屋に入れたい、と思う (欲求)

その欲求が他の欲求よりも強い

窓を開けるという行為がその効果をもたらす、と考える (信念)

その行為が自分の能力内にある、と考える

それ以外の行為は可能でなく、望ましくもない、と考える

そうならば、それ以外の事情が同じならば、

その欲求と (諸) 信念が、行為者に窓を開けるという行為をさせる。<sup>(36)</sup>

すなわち、なぜその人は窓を開けたのか、という問いに対して、それが意図的な行為である場合には、新鮮な空気を入れたくて、かつ、窓を開ければそうできると思ったからだ、という具合に、少なくとも彼自身の欲求と信念に言及して説明するのが普通である。この種のパターンの説明は、人間というものに関する我々自身の常識的な見解に深く根ざしており、人間の行為は今述べたような仕方で欲求と信念によって体系的に起きている、すなわち、行為者は合理的である、という基本的な考え方に基づいている。<sup>(37)</sup> これは一般に合理性の仮定と言われるものであるが、「合理性に関する寛容の原則」は、他人の行為を説明するときには、できるだけその行為者を合理的とするように説明せよ、というものである。

この原則は、「真理に関する寛容の原則」の場合と同様、探究のための経験的な仮説や方針ではないと考えられる。それを見るために、行為者が合理的であるというときに、我々がどんなことを意味しているかを考えてみよう。まず一つは、行為者の信念体系が合理的であるということであ

る。すなわち、その人の諸信念は整合的で一貫性がある、ということである。これをディヴィドソンは、「大体整合的で真なる信念体系<sup>(38)</sup>」を持つものとして話者を見ることができなければ、我々は話者を「合理的」とは見なせなくなると述べ、「帰属させられた欲求と信念は互いに適合しなければならぬ……という意味の合理性<sup>(39)</sup>」を行動の中に発見できなければ「目的論的説明」の説得力が得られない、と述べる。もちろん、整合性を厳格にとって、信念体系が論理的矛盾を全く含まないことを意味すると考えると、我々はだれもこれを満たすことができなくなってしまう。したがって、もう少しゆるく、一見して明らかな矛盾を含まない、たとえばどの信念も他の信念の否定になっていない、といった程度に考えるべきかもしれない。しかし、後で述べるように、信念と欲求による説明パターンにとって、寛容の原則は規範的な性格を持つものであり、現実の人間の不完全な推論能力や信念体系の不整合性を説明するためのものではないことに注意すべきである。その意味で、行為者に信念を帰属させるときは、その信念体系ができるだけ合理的となるようにせよ、というのが「信念の合理性に関する寛容の原則<sup>(40)</sup>」である。

行為者が合理的であると言うときに我々が意味するもう1つのことは、欲求と信念から行為が帰結する仕方が合理的である、ということである。ここでの合理性は、「説明されるべき行為は、帰属させられた欲求と信念に照らして理にかなって (reasonable) いなければならない、という意味<sup>(41)</sup>」である。そこで、行為を説明するときに、行為（不作為でも含む）を行なう理由として十分説得的となるように、行為者に欲求と信念を帰属させよ、というのが「行為の合理性に関する寛容の原則」である。人は完全に合理的であると仮定することはもちろんできないが、しかし少なくともかなりの期間にわたって相当程度にこのような合理性を欠いているなら、その主体に欲求や信念を帰属させての行為の説明は意味をなさない。そもそもその行動は、行為としての説明対象にならないと考えるべきである。合

理性は、欲求、信念、行動の結びつき具合をある種のパタンの中に位置づけて当の行動を見る（説明する）ことに存するのだから、あまりにそのパタンから逸脱することは、もはや行為としての要件を備えないことになる。

こうして、行為、欲求、信念、意図などの概念を用いて行なわれる行動の説明にとって、合理性に関する二つの寛容の原則、「信念体系の合理性」「行為の帰結の仕方の合理性」は、経験的仮説ではなく、説明そのものの成立にかかわる構成的原理であると考えられる。

信念と意味を帰属させて人の発話を解釈するときに、ある信念を帰属させるということがその発話のある仕方で解釈することであり、別の信念を帰属させることがすなわち別の仕方でその発話を解釈することであったが、ちょうどそれと同じように、行為者の心的態度を欲求と信念に分けて、ある欲求-信念を帰属させることがその人の行為のある仕方で説明することであり、別の欲求-信念を帰属させることはすなわち別の仕方でその人の行為を説明することである。意味と信念が解釈において互いに補完的であるのと同じく、欲求と信念（および他の諸態度）も、行為の説明においては互いに補完的、相互依存的であると考えられる。解釈の場合と同様、行為の説明のこのような全体論的な性格から、これらの寛容の原則の構成的原理という性格が由来すると思われる。

#### IV

さて、発話は一般に話者の信念をなんらかの形で表わすものであり、行動も一般に行為者の信念をなんらかの形で表わすものである。そして、発話は行動の一形態にほかならないのだから、一般に行動は信念の表われである。このような言い方は、きわめて常識的ではあるけれども、重大な誤解を招きやすい。それは、あたかも、人の心の中には(ある時点において)かなり定まった信念内容が確固として存在する、ということを示唆してい

るからである。その存在する信念が外に現われて来たものが行動だとすると、我々が行なおうとしている行為の説明において、探究の方向はなによりもその人の心の中へと向かわなくてはならなくなる。そこでは、彼が心の中に抱いている信念に、どのような経験的証拠を手がかりにしたら近づくことができるだろうかとか、それはどの程度まで可能だろうかといった問題が重要となる。しかし、我々はその前に、欲求と信念に言及して行なわれる説明の構造そのものにまず注意を向けるべきであろう。と同時に意味に言及して語られる「言語の意味」の説明の構造そのものにもまず注意を向けるべきであろう。<sup>(42)</sup> いったい、信念が我々の心の中に存在するというのはどういうことなのか。

ある人が片手を上げたとする。なぜ彼は手を上げたのだろうか。常識心理学は、友人の注意を引きたくて、かつ、手を上げれば友人の注意を引くだろうと思っていたから、と欲求と信念に言及して説明する。これは、手を上げるという比較的明らかなことを、それよりはるかに説明者にとってあやしいこと<sup>(43)</sup>（欲求と信念）を用いて説明している。この説明の有効性は、説明されている種類の出来事が、措定された欲求と信念によってさらに次から次へと説明されていくというところにある。<sup>(44)</sup> つまり、行為という種類の出来事が、目的論的説明によって、一貫性を持つものとして行動パタンの中へはめ込まれるところに意義があるのである。したがって、このパタンにはめ込まれるためには、行動はある種の規範の下で見られなければならない。この規範は、言語的には、欲求、信念、意味、意図、行為といった概念の適用に関する規範として現われる。

信念という概念が適用されるためには、……次の二つの条件が満たされなければならない。(1) 一般に、通常、どちらかと言えば、もし  $x$  が  $p$  と信じているならば、 $p$  は真である。(2) 一般に、通常、どちらかと言えば、もし  $x$  が  $p$  と言うならば、彼は  $p$  と信じている（そして

(1) より  $p$  は真である). これらの条件が満たされないと, 合理的な, コミュニケーションを行なうシステムはないことになってしまう. 信念を持つ者 (believers) や信念を表明する者 (belief-avowers) はいないことになる.<sup>(45)</sup>

これは, 寛容の原則のうち, 信念の真理性に関する寛容の原則と重なる. この前提が一般に満たされないと, 信念を帰属させての意味解釈や行為の説明はその基盤を失うことになる.

また, 行為の目的論的説明においても, 信念と欲求の帰属には, 次のような規範が課せられる.

行為を説明する信念と欲求は, その信念と欲求を持つ人ならだれでも, そのような仕方で行為する理由を持つであろう, というものでなければならない.<sup>(46)</sup>

つまり, 信念欲求が心の中にあってそれが行為として外に現われたという図式ではなく, 行為を信念・欲求という要因によって起こるものとして見るというのが, 我々の分析結果である. ある程度に組織化された常識が行なっているのは, まさにそのようなルールに従って行為者にあるパタンの信念・欲求を帰属させ, あらためて, その信念・欲求の結果としてその行為が行なわれたと言い直すのである. 「一般に, どの人の信念や欲求も, 状況が与えられたら, 彼が持つべき (ought to have) 信念や欲求である」<sup>(47)</sup>と認めるのである.

したがって, 欲求と行為の矛盾, 合理性の欠如, 偽なる信念, 誤りなどについては, 特別の説明が与えられなければならない.<sup>(48)</sup> のどが乾いていたのに彼はなぜそばのコップの水を飲まなかったのか. それが水でないと思ったからだ. いや, 水だとは思ったが, 汚いと思ったからだ. でも, あo状況で汚いと思う理由はないではないか. いや, その前にもあoコップが汚いことがあったからだ……. 我々は, 彼の行為 (不作為) が, 彼の信念

と欲求の下で合理的であれば、その限りでひとまず納得する。

とはいえ、我々は寛容の原則を構成的原理としてふりかざすだけでなく、その及ぶ範囲や強さをもっとていねいに規定する必要がある。また、この寛容の原則は、常識心理学と科学的心理学の境界線を描くものとして、その妥当性が議論されることがある<sup>(49)</sup>。さらに一般に、物理的なものを扱う理論と、心的なものを扱う理論と区別する特殊な原理としての性格を負わされることもある<sup>(50)</sup>。信念を措定する説明理論が、そもそもどのような効用を持つか、言語の意味や他者の理解にどう寄与するのか、といった問題にとって、寛容の原則の射程を見定めることはきわめて興味深いと言えよう。

#### 注

- (1) もちろん、現実に英語や日本語の辞書で伝統的に採用されている単語や接辞などが、理論的な観点からもその言語の最小要素とは限らない。また、ある自然言語に属する文の集合の濃度が可算無限であるのか、あるいはそれ以上であるのかについては議論がある。Langendoen and Postal (1984) 参照。しかし、ここでは生成文法の基本仮定に基づいて、可算無限であるとしておく。
- (2) 「雪は白い」と「雪が白い」は、総称性や焦点などの点で重要な相違があるが、それがどのように真理条件に反映するかは必ずしも明らかではなく、目下の問題とは別に十分な検討が必要である。ここでは、その差に立ち入らずに適宜使い分けることにする。
- (3) 翻訳理論および翻訳関係については、クワインの「翻訳の不確定性」という興味深いテーゼがあるが、本稿ではそれについては直接は扱わない。Quine (1960) 参照。
- (4) Davidson (1973), p. 134.
- (5) 「『雪は白い』は真である iff 雪は白い」が「『雪は白い』は雪は白いということの意味する」と同義だと言っているのではない。
- (6) Davidson (1973), p. 134, Davidson (1974), p. 150.
- (7) Davidson (1973), p. 128.
- (8) Davidson (1973), p. 139.

- (9) Davidson (1973), pp. 125-6.
- (10) Davidson (1974), p. 144.
- (11) Davidson (1974), p. 144, p. 152.
- (12) 後述のようにこれは *T*-文ではない.
- (13) Quine (1960), §§ 9-10.
- (14) Davidson (1975), p. 168.
- (15) Davidson (1974), p. 152.
- (16) Davidson (1973), p. 137.
- (17) IIIを参照.
- (18) Wilson (1959), p. 530.
- (19) Wilson (1959), p. 532.
- (20) Quine (1960), §§ 13-15.
- (21) Dennett (1971), Dennett (1972), Dennett (1981a), Dennett (1981b),  
Dennett (1981c), Dennett (1982), とりわけ簡潔なものとして Dennett  
(1987a).
- (22) Sober (1978).
- (23) Cohen (1979), Cohen (1981).
- (24) Stich (1984), Stich and Nisbett (1980).
- (25) March (1978), Black (1985).
- (26) Hempel (1962).
- (27) Devitt and Sterelny (1987).
- (28) Dennett (1971). また, 注 21 を参照.
- (29) Thagard (1983). また, Sesardić (1986).
- (30) Davidson (1975), p. 168.
- (31) Davidson (1975), p. 169.
- (32) Davidson (1974), p. 153.
- (33) Davidson (1973), p. 137.
- (34) ただし, Davidson (1974), p. 151.
- (35) Quine (1960).
- (36) Horgan and Woodward (1985), p. 197 を参照.
- (37) Horgan and Woodward (1985), p. 197.
- (38) Davidson (1973), p. 137.
- (39) Davidson (1975), p. 159.
- (40) この意味の合理性の程度について具体的に論じたものとして, Føllesdal



- (1981), Føllesdal (1982), Thagard (1983).
- (41) Davidson (1975), p. 159. しかし、素人にとってはなるほどもっともだと思われる推論や行為も、確率論や決定理論の立場からは誤まった直観にすぎないといわれることがたくさんある。とばく者の誤解がその代表例であろう。人間の非合理的な側面を強調して、合理性仮定に対する反例として利用されることが多いが、そのようなデータをどう評価すべきかは決して自明ではない。Stich (1984) を参照。
- (42) 黒田 (1971), p. 230.
- (43) Davidson (1975), p. 159.
- (44) Davidson (1975), p. 159.
- (45) Dennett (1971), p. 18.
- (46) Davidson (1975), p. 159.
- (47) Dennett (1971), p. 19.
- (48) この観点から寛容の原則を論じたものとして Collin (1987).
- (49) Cambell (1986), Devitt (1987), Root (1986).
- (50) Davidson (1973), Davidson (1975), Davidson (1976), Dennett (1971) および注 21 参照。

#### 参考文献

- Asher, N. (1985). "The Trouble with Extensional Semantics", *Philosophical Studies*, 47, pp. 1-14.
- Black, M. (1985). "Making Intelligent Choices, How useful is decision theory?", *Dialectica*, 39, pp. 19-34.
- Cambell, K. (1986). "Can Intuitive Psychology Survive the Growth of Neuroscience?", *Inquiry*, 29, pp. 143-152.
- Collin, F. (1987). "Meta-Constraints upon Interpretation", *American Philosophical Quarterly*, 24, pp. 137-148.
- Davidson, D. (1973). "Radical Interpretation", in Davidson (1984), pp. 125-140.
- (1974). "Belief and the Basis of Meaning", in Davidson (1984), pp. 141-154.
- (1975). "Thought and Talk", in Davidson (1984), pp. 155-170.
- (1976). "Reply to Foster", in Davidson (1984), pp. 171-179.
- (1982). "Empirical Content", in LePore (1986), pp. 320-332.

- (1984). *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- (1985). "Incoherence and Irrationality", *Dialectica*, 39, pp. 345-354.
- Dennett, D. (1971). "Intentional Systems", in Dennett (1978), pp. 3-22.
- (1972). "Reply to Arbib and Gunderson", in Dennett (1978), pp. 23-38.
- (1978). *Brainstorms*. Brighton: The Harvester Press.
- (1981a). "True Believers", in Dennett (1987b), pp. 13-42.
- (1981b). "Three Kinds of Intentional Psychology", in Dennett (1987b), pp. 43-82.
- (1981c). "Making Sense of Ourselves", in Dennett (1987b), pp. 83-116.
- (1982). "Beyond Belief", in Dennett (1987b), pp. 117-212.
- (1987a). "Mid-Term Examination: Compare and Contrast" in Dennett (1987b), pp. 339-350.
- (1987b). *The Intentional Stance*. Cambridge: The MIT Press.
- Devitt, M. and Sterelny, K. (1987). *Language and Reality*. Oxford: Basil Blackwell.
- Føllesdal, D. (1981). "Understanding and Rationality", in Parret (1981), pp. 154-168.
- (1982). "The Status of Rationality Assumptions in Interpretation and in the Explanation of Action", *Dialectica*, 36, pp. 301-316.
- Horgan, T. and Woodward, J. (1985). "Folk Psychology is Here to Stay", *Philosophical Review*, 94, pp. 197-226.
- Kornblith, H. (1985). *Naturalizing Epistemology*. Cambridge: The MIT Press.
- 黒田 亘 (1971). 「他人のころろ」, 『経験と言語』, 東京, 東京大学出版会 (1975) に再録, pp. 229-248.
- Langendoen, D. T. and Postal, P. M. (1983). *The Vastness of Natural Languages*. Oxford: Basil Blackwell.
- LePore, E. (1982). "In Defence of Davidson", *Linguistics and Philosophy*, 5, pp. 227-294.
- (1983). "The Concept of Meaning and its Role in Understanding Language", *Dialectica*, 37, pp. 133-139.
- (ed) (1986). *Truth and Interpretation: perspectives on the phi-*

- Philosophy of Donald Davidson*. Oxford: Basil Blackwell.
- March, J. G. (1978). "Bounded rationality, ambiguity, and the engineering of choice", *Bell Journal of Economics*, 9, pp. 587-608.
- Parret, H. (ed.) (1981). *Meaning and Understanding*.
- Quine, W. V. O. (1960). *Word and Object*. Cambridge: The MIT Press.
- Root, M. and Wallace, J. (1982). "Meaning and Interpretation", *Notre Dame Journal of Formal Logic*, 23, pp. 157-173.
- Stich, P. S. (1984). "Could Man Be an Irrational Animal? Some Notes on the Epistemology of Rationality", in Kornblith (1985), pp. 249-267.
- Stich, S. P. and Nisbett, R. E. (1980). "Justification and the Psychology of Human Reasoning", *Philosophy of Science*, 47, pp. 188-202.
- Sesardić, N. (1986). "Psychology without Principle of Charity", *Dialectica*, 40, pp. 229-240.
- Thagard, P. and Nisbett, R. E. (1983). "Rationality and Charity", *Philosophy of Science*, 50, pp. 250-267.
- Wilson, N. L. (1959). "Substances without Substrata", *Review of Metaphysics*, 12, pp. 521-539.