

文体論における内在的標準に関する小考

内 山 和 也

【要 旨】

近代的な文体の概念は、その認識の端緒において文脈と密接に関わっている。しかし、文脈の概念は複雑かつ多義的であり、文体論には文脈の概念をさまざまな観点で検討することが求められる。本稿では、主に文字列における文脈の概念を検討し、テキスト生成AIとの類推によっても文脈上の対比について考察した。その結果、文脈から文体の認識や概念を導くことが人間を言語表現の受け手として特徴づけるものであると述べた。

【キーワード】

日本語文体論 文脈 対比 文字列 テキスト生成AI

0. はじめに

文体論の研究史は、文体概念、すなわち「文体とは何か？」との格闘の歴史であると言われることがある。言い換えれば、いかにすれば言語的な直観を言語形態（言語の表層の秩序）によって裏付けることができるのかという問いを繰り返してきたということでもある。

たとえば、リファテール（1978）は、言語表現に意味を損ねることなく付加された情報、すなわち余剰なメッセージが文体であると考えている。この考え方によれば、文体は本質的に不要なものでありながら、あらゆる言語表現が不可避に担うべきものでもある。なんとなれば、まったく特徴のない言語表現であってさえ、特徴のないこと自体が文体と見做されうるからである。この意味での文体は、読み手による読みの中にその手がかりが求められることになる。すなわち、読みの中に感じとられる違和感が文体のトリガーだと考えられ、比喩的には読みにおける読み手の反応速度の遅延（リファテール1978）、あるいは「表現の手触り」として捉えることができる。言語表現の表面を指先でなぞったとき、どこかで引っかかってしまうことなく動きながらも指先に何か感じるものがあるといったイメージである。このように文体の認識の端緒とは、前後のつながりのズレや他処との流れの違いといった、文脈上の対比によって生じると考えることができるのである。

この見方は、一方で、文体を逸脱 deviation と見做す立場とも重なり合う。逸脱の文体論とも呼ばれるこの立場は、ある規範 norm と比較してどのような言語的な特徴の違いが見られるかを問うもので、その規範を統計的な平均値などのテキスト外に求める場合（外在的標準 external deviation）とテキスト内の他の部分に求める場合（内在的標準 internal deviation）とがある（Levin 1965）。後者は、言語表現の前後関係から予期される表現に対しての逸脱であって、裏切られた期待 defeated expectancy という言い方がされることもある（Leech1969：119f.）。

文体を余剰な付加情報（による表現の彫琢や前景化の原理）と見る場合も、内在的標準からの逸脱（による表現の彫琢や前景化の原理）と見る場合も、文体の認識は言語表現における前後関係との関係性が端緒と見做されるという共通点が認められる。言い換えれば、文体あるいは文体の認識は、文脈と密接な関わりがあるということである。しかしながら、一方で文体という概念が実に定義の難しいものであるように、一方で文脈もその意味するところは多様でありうる。本稿では、書きことばの表現における文脈、特に読みの文脈について、文体論の視点で検討したいと考える。

1. 文脈の概念の多義性

言語表現において、文体との関係でしばしば用いられる語に「文脈」がある。周知の通り、文脈の概念は言語研究において多義的である。

まず、一般に文脈は、言語的な（あるいは言語内の）文脈と非言語的な（あるいは言語外の）文脈とがあると見做され^(注1)、後者は、さらに、物理的な環境に属する文脈と世界^(注2)の知識に属する文脈とに区別される。

言語記号の根底的な原理として、記号の恣意性と話線の線条性(展開の1次元性)とがあるが、後者から言語表現は必ず前後関係をもって生じるという事実が導かれる。このとき、話線に実現したそれぞれの語や文などの単位について見れば、それが表わす内容は前後関係によって理解されると言える。例えば、「真っ赤なシャツを着る。」の「真っ赤な」は色彩を表わすが「真っ赤なウソをつく。」の「真っ赤な」は色彩を表わさないといったことである。かかる言語表現の理解に関わる言語内の情報が、言語的な文脈と呼ばれるものである。空間的・時間的性質として、ある言語単位や言語要素と近接した位置にあるか生じるかする言語単位や言語要素の総体を指しており、特定の言語単位や言語要素における前後関係は、環境 environment と呼ばれることもある。

また、表現学では、言語表現の特徴のひとつとして、言語表現が始められる（開始できる）こと、積み重ねられる（連続してつながられる）こと、閉じられる（終了できる）ことを挙げる^(注3)。これらのことを前提に文脈というときには、文や文章での先頭（の要素）から終結（の要素）に至る流れやつながりの有無または種類を指すことになる。この場合の文脈は、脈絡と言われることも多い。

さらに、文章表現の史的な跡付においては、「漢文脈」や「欧文脈」などの術語も用いられる。これらは和文（平安和文を範とする文章）を基調とする文章表現（これを和文脈とする）のなかに異質性を感じさせるものとして流れ込んだ文章表現上の要素をいう（木坂1976；1988）。和文脈に流入した漢文（変体漢文）と漢文訓読およびそれらに準じた文章表現が漢文脈である。同じく、欧文の語法や文法・文型などの特徴を模した表現（欧文直訳調のいわゆるバタ臭い表現）が流入したものを欧文脈と呼ぶ。ここでは、文章表現を均質なものとして統合する働きが文脈と呼ばれているもので、逆に異質性を感じさせる文脈（混入した文脈）は、文章脈という術語で区別されることもある。文脈と文脈の対比によって文脈の異質性（＝文章脈）の認識が導かれるのである。

一方、非言語的で物理的な意味での文脈は、状況 situation（場面、場とも訳される）とも呼ばれる。この意味での文脈は、話し手の置かれたリアルな環境や実際の場面（話し手と聞き手の関係性といった社会的要因もここに含まれる）などの言語使用上の諸条件を指し、そのため、話しことば（対面での会話）に代表されるリアルタイムでの言語使用あるいは言語行動について

言われることが多くなる。したがって、物理的な文脈は、言語（表現）を言語外の現実結びつける中間的な媒介項であると見ることができる。

また、非言語的だが物理的でない、共通知識 common knowledge という意味での文脈は、言語的なメッセージの伝達に参与する言語外的な知識のことである。もっとも言語内の知識と言語外の知識とを明確に境界できるのかという議論はありうるだろうが、この意味での文脈は『標準的な人間』が特定の言語においてどのようなものであるのかを推定させるものである。なお、ここで「特定の言語において」ということわりが必要なのは、社会的・文化的な環境や背景を共有することによってこそ、『標準的な人間』が持つべき知識ないしはデータ^{注4}、行なうべき推論が一定の共通性を有すると前提されることになるためである。

1.1. 逸脱的なものとその解釈

文体（スタイル）style という語は、粘土板に文字を刻むための道具であった尖筆（スタイラス stylus）に由来する。語源という意味では、書き手の個性性としての文体の認識は、理論上、文字の発生にまで遡ることができるのである。現在のところ知られている最古の文字は、紀元前3000年紀にメソポタミアで用いられたシュメール象形文字であり、同じく古代メソポタミアに遡ることができるものに（西洋）占星術がある。バイジェント（2021：67f.）によれば、占星術とは、天空の星々の配置などの自然現象を神のメッセージと見做し、そこから異例の現象の前兆を読みとることで、ありうる未来を知ろうとするものであるという。その結果、観測を通じて得られた天体の運航の規則性によって、占星術は天文学だけでなく統計や確率といった数学とも接近していくことになる。

当然ながら文体論と占星術には、いずれも古代メソポタミアに遡りうるという点以外に何ら関連性はない。一方で、両者の考え方には一定の共通性も指摘できるように思われる。つまり、メッセージを均質的な流れを有するものとし、そこに見られる逸脱的なものをトリガーとして、そこに何らかの解釈を施すことで逸脱を理由づけようとすることは、人間の原初的な思考のパターンのように思われるのである。

1.2. 文脈と文体論

前節までの検討から、文脈は(1)言語表現の時間的な布置、(2)言語使用の物理的な条件、(3)書き手と読み手の共通知識という3つの視点で捉えられるものであることがわかる。また、これらの視点は互いに排除し合うものであるとは言えないため、文脈の捉えられ方は複合的なものになることも少なくない。このように複合的な多義性を有しながらも、言語表現について文脈が言及されるのは、言語表現が担う意味あるいは言語的メッセージの解釈が文脈によって限定されることが明らかであるからだろう。これは、生物の行動一般が環境要因に制約されるのと同様であり、読み手の位置に人間を想定するかぎり、疑いをいれることは難しいものと考えられる。

その一方で、文脈には大きな課題も指摘される。つまり、言語表現あるいは言語的なメッセージの解釈が文脈に限定されるのだとしても、特定の解釈と対応する文脈の条件を限定する定式化は困難だという点である（カラー1985：197ff.）。そのため、言語表現に関して文脈を取り扱う際には、その論点や問題設定において特に重要性が高いと思われる条件のみを選択的にとりあげることになる。しかし、文脈は複合的な多義性を有し、種々の要素や要因が絡まりあっているため、条件の選択は必ずしも容易ではない。研究の核心が文体の概念そのものに置かれてきた文体論においては、文体を付加情報や逸脱と考えれば言語表現の時間的な布置としての文脈の問題とは切り離せず、あるいは、文体を表現形式の選択に関わるものと考えれば言語使用の物理的な条

件としての文脈の問題に通じ、直喩・隠喩・換喩などの比喩表現の使用を文体上の特徴と見做すのならば書き手と読み手の共通知識としての文脈の問題とも関わりあわなければならない。ムーンマン（1970：219）が指摘するように「文体は人間的な現象で非常に複雑性を持つ……いつの日にか文体についての説明をまとめ上げられるとすれば、その説明は極度に複雑なものであろう」と考えられるのだとすれば、文脈についての理解もその大きな一部をなすことになるだろうし、そのために、文体論では文脈の概念をさまざまな点から検討する必要があるであろう。

2. 文字列における文脈

第1言語であるか第2言語であるかを問わず、文字・表記の学習においては、文字の流用あるいは字体の代用が生じることがある。古くは、土居（1955）において、欧州での日本語教育で学習者がアルファベットのようにカナ文字を記す様子が報告されている。また、内山（2015；2016）では、日本語の上級レベルの学習者にあっても、カタカナの「ホ」が漢字の「木」によって、ひらがなの「し」とカタカナの「レ」が互いに、ひらがなの「る」が漢字の「子」の崩し字によって代用されている例が見られるとする。これらはいずれも第2言語学習（非日本語母語話者による日本語学習）での事例であるが、第1言語（日本語母語話者による日本語学習＝国語学習）であっても、たとえば、小学校で漢字を学んだ児童が「急」という漢字を「ク」（カタカナ）・「ヨ」（カタカナ）・「心」（漢字）と分解して覚えようとするものがあつたとすれば、これも文字の流用のひとつといえることができるであろう。語学の学習においては、学習者は自らが利用することのできるリソースを習得の目標とする言語に差しむけること（例えば、韓国語での後置詞の知識を日本語の格助詞の習得に活用したり、中国語での漢字の知識を日本語の語彙の習得や読解に活用するなど）ができる以上、文字についての既存の知識である字体を活用することも許されるのである。

そもそも文字では、それらが形象としては同等であっても、互いに異なる文字として認識されることがありうる。たとえば、漢字の「口」とカタカナの「口」はストローク（点画）の組み合わせ（丨・フ・一の3つのストロークが接合されて方形を成す）という点で見れば同等である。活字やデジタルフォントでは、漢字である「口」はカタカナである「口」よりも大きめに実現・表示されるが、これは純粋にデザイン上のものであって、ストローク（点画）の違いによるものではない。一般に、ストローク（点画）の基本的な組み合わせとしての文字の構造的な原型は字体と呼ばれ、字体に施されたデザインや意匠は書体、印刷されたりディスプレイ上に表示されたりした個々の文字の具体的な形は字形と呼ばれる。先ほどの例についていえば、漢字の「口」とカタカナの「口」とは、書体における字形の違いは認められても、字体においては同等であると言える。では、字体における両者の違いがどのように識別されるのかといえば、それぞれの字体がどのような文字体系に属するのかによる。文字論では、「あ」や「ア」のような個別のものを文字と呼ぶのに対して、ひらがなやカタカナといった集合的なものを文字体系と呼んで区別する。この視点からは、それらが字体として同等であっても、漢字という文字体系に属する「口」の字体と、カタカナという文字体系に属する「口」の字体とは、異なる文字の字体として区別されるのである。しかし、この区別は理論的なものにすぎないため、実際の表記（を読む・読みとるとき）において両者がどのように区別されているのかは、別の観点から説明されなければならないだろう。そこでとりあげられるのが、文脈という概念である。大雑把に言えば、「出入口」のようにほかの漢字と並んで用いられていれば漢字の「口」として読みとられ、「セイロ」のようにほかのカタカナと並んで用いられていればカタカナの「口」として読みとられことになるだ

ろうということである。文字列の前後関係の中で先行する要素や後続する要素と結びつけて、いわゆる語を構成するように解釈されるわけである。したがって、そのような読みの文脈がなければ、同等の字体のいずれが漢字の「口」で、いずれがカタカナの「口」であるのか識別することは、実践的にできないのである。

ひとくちに漢字の「口」といっても、その物理的な実現である字形は、千差万別である。手書きの文字であれば、仮に同じ人物によって書かれたものであっても、字形ごとには必ず違いが見られるのであり、印刷された文字でも細部まで字形が一致するというわけではない。また、電子テキストにおいても、電子ファイルのテキストデータ内に格納された文字コードという点では同じということができるが、それがディスプレイ上に実現されるときには何らかの物理的な異なりが見られるはずである。そうだとすれば、読み手は、そのような限りもない字形の変異の中から特定の字体を抽出し、その字体から文字を識別し、さらに、その文字が構成する語などの言語単位を理解していることになる。すでに漢字の「口」とカタカナの「口」の例で見たように、字体から文字を識別するとき、文脈が関与していることは明らかである。しかし、「避難は非常口から」という文字列の中で「口」の字体が漢字の「口」（の字体）であると識別されるためには、「非常口」という文字列が特定の語を表わすと理解されなければならず、文字列から「非常口」という語が取り出されるためには、あらかじめ「口」が漢字の「口」（の字体）であると識別されていなければならない。つまり、文脈に沿った読みとりは〈字形→字体→文字→語〉のように線的になされるのではなく、前後関係が常に総体として解釈されていることになる。

マルティネ編（1972：371）は、言語記号は形式と内容とに区別できるため、言語記号の形式の面（音素や形態素）に影響する文脈と言語記号の内容の面（概念や意味）に影響する文脈とに区別することができるとする。また、形式の面に影響する文脈は、音素（の変異）に関わる文脈と形態素（の変異）に関わる文脈があるとする。しかし、このような区別は、言語を本質的に音声的なものであると見做す場合に成り立つものであり、字形・字体・文字に関わる文脈が形式の面に影響する文脈に含まれるのか定かではない。

このことから知られるように、言語的な文脈の問題は、文字・表記に限らず広く日本語の書きことばの表現について考えるときにも、なお検討が必要であるように思われる。つまり、文脈という語は本質的に多義的でありながら、特に日本語のように複雑な表記を採用する言語にあっては、文字列における働きが十分に明らかであるとは思われないからである。なお、表記の複雑性と文脈の関係については、別稿を用意したいと考える。

3. 新たな文脈の機能

3.1. テキスト生成AIにおける文脈

言語的な文脈は、異なる観点から見ることでもできる。2023年になって巷間を賑わせた ChatGPT（チャット GPT）に代表されるテキスト生成AIについての知見である。

テキスト生成AIの衝撃は、その汎用性や対話によるインタフェース、テキストの流暢さや多言語への対応、無料で利用できるバージョンの提供といったことが複合的に含まれていたと思われるが、岡嶋（2023）は文体や語調が極めて人間らしいことが大きく影響しているものとも指摘している。一方で、技術的に貢献しているのは、プロンプト Prompt（指示文）と呼ばれる手法であろう。生成AIに利用されている汎用の言語モデルは、大規模言語モデル（Large Language Models：LLM）と呼ばれる確率モデルであり、膨大な学習データから確率計算による分類問題^{註5}の逐次遂行を繰り返すという仕組みになっている。

テキスト生成AIにおけるプロンプトは、確率計算の文脈を与えることに相当するものとされる（鈴木2023）。まず、プロンプトによる質問の入力によって確率計算が遂行され、テキストが生成される。次に、AIが生成したテキストに対し、さらなるプロンプトによる質問が入力されることで、それが新しい文脈となって、そこからまた確率計算とテキストの生成が続いていくといったことである。このようなプロンプトによるタスクの実行は、質問の内容（何をどのように入力したか）に依存するもので、質問者（誰が入力したか）には依存しない。そのため、テキスト生成AIでは、どれだけよい質問ができるのかという意味で、ユーザの側の表現力が回答の有効性に関わっていることになる。このことから、有効なプロンプトをいかにして見つけ出すかという試行錯誤、いわゆるプロンプトエンジニアリング（プロンプト工学）の必要性が言われることになる。

ここでの文脈とは、言語的に見れば、言語表現の時間的な布置を意味することになるだろう。これまでもウェブ上のテキストが新たなデータを提示することで、文章表現や文字表記の研究に進展を与えたということが言えようが、テキスト生成AIによって、定式化が難しいと考えられてきた文脈という概念をよりよく理解できる可能性があるのではないだろうか。

3.2. 人間とAIにおける表現の手触り

すでに述べたように、文体を付加情報あるいは内在的標準からの逸脱と見るとき、文体の認識の端緒となる文脈上の対比（言語表現における前後関係との関係性）は「表現の手触り」という比喩で言い表わされることがある。では、人間が感じるように、テキスト生成AIが「表現の手触り」を感じることはあるのであろうか。

3.2.1. 人工知能と人間の知性

まず、検討しなければならないのは、AI（人工知能）に知性はあるかという観点であろう。

タクル（1996）は、現代的な思考の特徴は、並立定義と視点循環にあると指摘している。これは一見矛盾した定義が混同されることなく両立するということであり、そのようにして判断や物事の境界を曖昧にすることがテクノロジーの特徴であるとも考えられる。たとえば、ロボットベットの接し方で、仕事をするときに邪魔だからと電源を切る行動をする人が、それが故障するとロボットが死んでしまうと取り乱すといった場合である。AIについても、同様の循環があると考えることができるだろう。テキスト生成AIについて言えば、実際の処理は確率計算を繰り返しているだけであり、なにひとつ理解しているわけではないと知りながら、テキストでの対話が成り立つと感じたときにAIが言語だけでなくさまざまなことを理解しているかのようにも感じるということである。そのときには、AIには知性があるけれどもなく、知性がないけれどもあるといった並立と循環とが成り立っていることになろう。

このことは他者から見た知性の有無であるが、ネーゲル（1989）は「コウモリであるとはどのようなことか？ What is it like to be a bat?」という問いを立て、コウモリによる内部の視点なしにはコウモリの体験の十分に客観的な記述はできないという「経験の主観的性格」を主張している。これにしたがえば、人間の主観は決して人工物には再現（シミュレーション）できないことになる。一方で、西田（2023）は、コンピュータの能力が飛躍的に向上して人間の主観が模倣（エミュレーション）できたとすれば、人間の意識は肉体が消滅した後も生き残ることになり、死の超越という既存の社会の前提を根本から覆す事態が起ころうと指摘する。ここでは後段の推測は措くとしても、コンピュータの演算能力の指数関数的な向上を考えれば、人間の主観を模倣的に動作（エミュレート）することで内部の視点から人間の主観が再現される可能性はありうると

見做すべきなのではないかだろうか。

3.2.2. 自然言語理解と確率計算

続いて検討すべきは、AI（人工知能）は自然言語の意味を理解しているのかという問題である。

自然言語理解は、かつてAIコンプリート（AI完全）とも言われた難問であり、それが達成できればシンギュラリティ singularity（人間の知性をAIが超えてゆく技術的特異点）となるともされてきた。実際に人間が自然言語をどのように理解しているのかということは完全には解明されていないが、膨大なデータと確率計算によるAIとは違う方法である可能性が高いと思われる。その一方で、AIが自然言語を人間と同じレベルで理解しているのかという問いに対しては、その可能性があるかと肯定的に評価する研究者も少なくないようである。もちろん、この場合の「人間」というのが誰を指すのか、それが特定の個人でないとなれば人間という一般的で普遍的な概念をいかに定義するのか、さらに同じレベルであることをいかなる基準で測定すればいいのかなど、さまざまな条件が必要とされることは当然である。

自然言語理解を問題にするときに関わってくるのが、物理的な身体のあることが認知機能にどのような影響を及ぼすのかという「身体性」embodimentの議論である。ホーキンス（2005）は、知能とは物事を記憶して類推によって予測する能力のことだと規定し、知能を備えた機械が人間に似ている必要はないと論じている。これまで知能に関して議論されてきた問題のひとつに、記号接地問題 symbol grounding problem と呼ばれるものがある。言語の意味や意図を『真に』理解するには記号処理だけでは不十分ではないかという考え方につながるもので、AIが人間と同じ知能を持ちうるのであれば、記号（シンボル）を処理する言語モデルだけでなく、記号と現実の事物を身体によって結合する世界モデルが必要だと考えられることが多かった。しかし、テキスト生成AIがその可能性を示しているように、身体性すら確率計算によって言語的に表示できるのであれば、言語モデルだけで充分だという考え方もできるはずである。言い換えれば、物理的な身体で言語を説明するよりも、言語で物理的な身体を説明すべきだということなのである。

ライバー（1994）は、研究の方法には、対象を実際に作ってみせる方法と対象を理論的に説明する方法があるとし、特に文法の研究は、それによって文と同等のものを実際に作ることになるという点で、そもそも工学的なアプローチであると指摘している。現在の理論言語学の主流である生成文法は、この工学的なアプローチによるものであり、生成文法の論者たちはテキスト生成AIとは競合的な立場をとっているようである（たとえば、酒井編著（2022）など）。一方で、生成文法とは異なる立場を採用する認知言語学の論者は、言語を理論的に説明しようとするものであり、テキスト生成AIにも比較的好意的なようである（たとえば、町田（2023）など）。しかし、認知言語学の知見が、AIが実際に作って示したものと共通するのであれば、後者の方がより明快で説得的であり、認知をベースにした理論的説明の必然性や必要性は低くなるというべきなのではあるまいか。

実際には、テキスト生成AIは言語の意味も概念を理解せず、単に確率計算を行なっているだけである。言い換えれば、意味を捨てて確率で処理をする道を選んだということになる（岡野原2023）が、そもそも『意味』を心理学の問題に過ぎないとして言語学から退けたのはアメリカ構造言語学の立場であった。このアメリカ構造言語学を批判して現れたのが生成文法（変形言語学）であると考えれば、生成文法とテキスト生成AIとがともに工学的なアプローチを採用しながら競合関係になるというのはむしろ自然なのであろう。

3.2.3. 言語表現への社会的な許容

岡田(1995)は「口ごもるコンピュータ」という例を挙げ、身体性(物理的な身体を持ち、そのことで何らかの影響を受けること)は人工物が社会に馴染むために必要とされるものではないかと論じている。人間と機械との共存は全面的に社会の側の許容に依存しており、たとえばコンピュータであれば、それが人間らしい身体的な動作を伴うことが、社会の許容度を高めるために必要なのではないかとということである。しかし、テキスト生成AIが自然な対話(のシミュレーション)をするようになった現在では、この許容度の問題は、人工物がいかに人間らしい振る舞いをするかということよりも、人工物が人間を脅かさないものであるのかという方へ移っているように思われる。従来危惧されてきたのが、機械が自律的に動いて人間がそのスイッチを切れなくなるということであったとすれば、現在危惧されているのは、人工物が社会を混乱させるきっかけになってしまうのではというより現実的な問題であろう。

デリダ(1978)は、テキストの外部はないと指摘する。このフレーズが意味しうところはさまざまであろうが、ひとつにはテキストにメタレベルがないと解釈することができよう。テキストに関してあるいはテキストに対して言及するものは、それもまたテキストであり、言語に言及するメタ言語のような審級があるのではない。その意味では、テキスト生成AIが言語表現を生み出すということは、私たち人間が経験するテキスト(の空間)の中に人工物によるものがとりいれられるということである。もちろん、このことは単に量的な問題に過ぎないと見做すこともできる。人工物がいかなる言語表現を生成したとしたところで、テキストに人間が参加できなくなるわけではまったくないからである。その反面、テキスト生成AIが非常に不適切な内容の表現を生成した場合、それが誰の責任になるのかといった議論の余地は小さくない。現在普及しつつある主要なテキスト生成AIのサービスは、モラルや倫理の面で十二分な事前調整が試みられており、それがある程度うまくいっていることが、技術の普及の下地となっていると言えるだろう。しかし、そのような調整を経て普及する技術なのであれば、テキスト生成AIには社会を混乱に陥れるだけの潜在的な脅威があるとする意見も故なしとしない。

ここで3.2.で掲げた問題に立ち返りたい。すでに述べたように、テキスト生成AIが言語表現に及ぼす影響は、量的なものに過ぎないとも見られる。これは、現状のテキスト生成AIは、AI同士で対話をするというわけではなく、あくまで受け手に人間を想定していると思われるからである。したがって、対話型AIという言い方をしたところで、実際に対話しているのは人間の側だけである。その意味では、人工物が表現の手触りを感じることはおそくないであろう。しかし、それは身体を持たず言語表現の受け手にもならない人工物には、その必要性がないということではかない。言い換えれば、私たち人間にとっての身体や意味といったものが、そもそも余分で過剰なものだということになるのだろう。

冒頭で述べたように、文体は言語表現における余剰なメッセージであると見做されることがあり、その認識は文脈上の対比、すなわち、比喩的に「表現の手触り」とされるものを端緒にするものであった。あるいは、同じ原理から、文体は規範としての内的標準からの逸脱と見ることもできるものである。そうであれば、文脈を活用して対話をシミュレーションするテキスト生成AIは、文脈によって与えられる文体の認識を持たないのであろうか。別の言い方をすれば、表現の手触りを感じず、文体といった複雑極まりない問題にも踏み込まない人工物に対して、身体や意味など本質的に余剰なものを抱える人間の側が物足りなさのようなものを感じることはないであろうか。この問いについて、本稿では、文脈を文体と関係づけることが言語表現における社会の側の許容度に関わるものと考えたい。このことは、文体論の新たな検討課題になりうるものであろう。

4. まとめ

ここまで、文体論の視点から、言語的な文脈、主に読みの文脈について論じてきた。

文体の近代的な定義として、文体を余剰なメッセージであると見る見方も文体を内的標準からの逸脱であるとする見方も、読みにおける文脈の上での対比にその認識の端緒が求められるという視点を共有していると考えられる。そもそも文体認識の起源に立ち返って考えれば、メッセージの均質な流れを阻害するものを認め、それを解釈によって意味づけようとすることは人間の原初的な思考であるとも思われる。

文体の認識が文脈と密接に関わるとしても、文脈の概念そのものが複雑で多義的である。そのため、文体論は文脈の概念をさまざまな観点で検討するべきであろう。

まず、欧米の言語研究で十分に考慮されてこなかったと思われるものに、文字列と文脈との関わりが挙げられる。日本語のように複雑な表記のシステムを採用する場合、文字列において字形・字体・文字と文脈との単純でない関係性が観察されるが、それが十分に理論化されているとは言えない面も見られる。

また、極めて人間らしい言語表現を可能にしている現在のテキスト生成AIは、対話型のインタフェースにプロンプトと呼ばれる手法を利用している。これは質問文の入力によって確率計算に文脈を与えることに相当するものと考えられることができる。そのため、テキスト生成AIによる知見を通じ、文脈の概念をよりよく理解できるようになる可能性がある。本稿では、テキスト生成AIの現状からの類推によって、文脈の上での対比から文体の認識や文体の概念を導くことが人間を言語表現の受け手として特徴づけることになるのではないかと考えた。

しかし、文体や文脈の概念は複雑で多義的であり、AIの研究もまさに日進月歩であるので、本稿で導いた含意の有効性についても、なお検証が必要となろう。

付記

本稿の内容には、2023年10月28日に別府大学メディア教育・研究センター(37号館)4Fメディアホールで開催された別府大学大学院主催の講演会・シンポジウム「どうするAI—その可能性とこれからの社会」において「人工知能は言語表現の手触りを感じるか?」のタイトルで報告(口頭発表)した内容を含んでいる。なお、当該シンポジウムは、2024年1月現在、別府大学・別府大学短期大学部オープンエデュケーションルームにおいて動画が公開されている(<https://www.beppu-u.ac.jp/oer/ocw/lecture-meeting/ai.html>)。

注

- 1) たとえば、マルティネ編(1972:369f.)では、言語研究が生理学的・社会的・心理学的条件から総体的に独立するためには、言語的な文脈と非言語的な文脈(=状況)とを峻別するべきであるとされている。
- 2) ここでは、人間一般が知りうる範囲のことを指す。
- 3) これらのトピックは表現学会の学会誌『表現研究』の表紙裏面にも示されている。
- 4) ここで知識とデータを区別するのは、個人の経験や体験の集積や履歴などの体系的でないものをデータの側も含めるためである。
- 5) 分類問題 Classification Problem とは、機械学習の用語でデータをどのカテゴリー(離散的データ)に分類するか予測する課題を指す。

参考文献

- 内山和也 (2015) 「日本語学習者におけるひらがなのストロークについて」, 『2015応用日語學術研討會論文集』, pp. 1-20, 育達科技大學應用日語系.
- (2016), 「日本語学習者に見るカタカナの字形とストロークについて」, 『2016応用日語國際學術研討會論文集』, pp. 1-13, 育達科技大學應用日語系.
- 岡嶋裕史 (2023) 『ChatGPTの全貌：何がすごくて、何が危険なのか?』 (光文社新書1267) 光文社.
- 岡田美智男 (1995) 『口ごもるコンピュータ』 (情報処理学会編・情報フロンティアシリーズ9) 共立出版.
- 岡野原大輔 (2023) 『大規模言語モデルは新たな知能か：ChatGPTが変えた世界』 (岩波科学ライブラリー319) 岩波書店.
- カラー, ジョナサン (1985) 『ディコンストラクション』 (富山太佳夫・折島正司訳) [原書刊1982] 岩波書店.
- 木坂基 (1976) 『近代文章の成立に関する基礎的研究』 風間書房.
- (1988) 『近代文章成立の諸相』 和泉書院.
- 酒井邦嘉編著 (2022) 『脳とAI：言語と思考へのアプローチ』 (中公選書125) 中央公論新社.
- 鈴木潤 (2023) 「ChatGPT (対話型文章生成 AI) を支える技術」 (2023年6月2日国立情報学研究所オープンハウス2023基調講演).
- タークル, シェリー (1996) 『接続された心：インターネット時代のアイデンティティ』 (日暮雅通訳) [原書刊1995] 早川書房.
- デリダ, ジャック (1978) 『根源の彼方に：グラマトロジーについて』 (足立和浩訳) [原書刊1967] 現代思潮社.
- 土居光知 (1955) 「ひらがなとカタカナ」, 心編集委員会編『心』8 (7), pp. 55-59, 平凡社.
- 西田宗千佳 (2023) 『生成AIの核心：「新しい知」といかに向き合うか』 (NHK出版新書705) NHK出版.
- ネーゲル, トマス (1989) 『コウモリであるとはどのようなことか』 (永井均訳) [原書刊1979] 勁草書房.
- ベイジェント, マイケル (2021) 『古代メソポタミア占星術：前兆の科学と天空の知識』 (倉本和朋訳) [原書刊2015] 太玄社.
- ホーキンス, ジェフほか (2005) 『考える脳 考えるコンピューター』 (伊藤文英訳) [原書刊2004] ランダムハウス講談社.
- 町田章 (2023) 『AI時代に言語学の存在の意味はあるのか?』 ひつじ書房.
- マルティネ, アンドレ編 (1972) 『言語学事典：現代言語学基本概念51章』 (伊藤英ほか訳) [原書刊1969] 大修館書店.
- ムーナン, ジョルジュ (1970) 『言語学とは何か』 (福井芳男ほか訳) [原書刊1968] 大修館書店.
- ライバー, ジャスティン (1994) 『認知科学への招待：チューリングとワイトゲンシュタインを道しるべに』 (今井邦彦訳) [原書刊1991] 新曜社.
- リファテール, ミカエル (1978) 『文体論序説』 (福井芳男ほか訳) [原書刊1971] 朝日出版社.
- Leech, G. N. (1969). *A Linguistic Guide to English Poetry*. London: Longman.
- Levin, S. R. (1965). Internal and External Deviation in Poetry. *Word*, 21 (2), 225-237. Retrieved from <https://doi.org/10.1080/00437956.1965.11435425>