

言語情報の前方誤り訂正における歴史的アプローチ

人文科学の人類共有知への貢献

得丸 久文

概要: ヒトの言語と知能はデジタル進化している。言語コミュニケーション物理層のデジタル信号が①音節・②文字・③電気信号と進化したことにより、①無限の語彙、②知能の連続的発展である文明、③検索と電送が可能になった。インターネットはキーワード検索によって、「きわめて大量だが、信頼性の保証のない学際的な言語情報」を提供する。この情報を活用するうえで、「前方誤り訂正」は不可欠であり、書誌学と歴史学が大いに貢献することになる。

キーワード: デジタル言語学, 前方誤り訂正, 誤り訂正符号, 群, フラクタル,

Historical Approach in Forward Error Correction of Linguistic Information

How Humanities Should Contribute to the Collective Human Intelligence?

Kumon Tokumaru

Abstract.

Human language and intelligence are evolving digitally. Digital signals in physical layers evolved from (1) syllables, (2) a character set, and (3) electronic signals, and (1) infinite number of word signs, (2) serial development of intelligence in civilization, and (3) keyword search and electric transfer were achieved. The internet provides “voluminous interdisciplinary linguistic information with uncertain reliability”. By using such information, Forward Error Correction should be implemented with the help of Bibliography and History.

Keywords: Digital Linguistics, Forward Error Correction, Error Correction Code, Group, Fractal

1. はじめに: 言語と知能の三段階デジタル進化

(1) 最古のオトガイの化石が発掘された洞窟

「デジタル言語学」は、筆者が 2007 年に南アフリカのクラシーズ河口洞窟を訪問して生まれた。ここは最古の現生人類洞窟であり、現生人類だけがもつオトガイ(mandible)の最古の化石が見つかった場所である。ここで言語が生まれたに違いない。

に穿たれた巨大で安全な海蝕洞窟に住むようになると、新生児が誕生後 1 年間歩かなくなり、揺り籠の中であやされて過ごすようになる。大型霊長類は、一度の妊娠で一匹しか産まず、妊娠期間も長く、胎内で神経の髄鞘化が行われるため、生後すぐに母親にしがみついて移動できる早成動物である。敵や捕食者や虫に襲われることのない安全な住居を確保したヒトだけが、知能発達のために新生児期を過ごす「二次的晩成動物」に進化したのだ。[1]

寝たきりの新生児を保護するために、閉経後のメスが家族の一員となり、家族が三世代で構成されるようになって、共同体への利他的献身を特徴とする真社会性化した。ウィルソンは、真社会性の生活を続けるうちに、さらに真社会性が進み、共同体に独自の匂いや声を生み出すという。

[2] 13 万年前から 7 万年前の期間に、群れごとに独自の鳴き声をつくりだしていたと思われる。それが、7 万 2 千年前にインドネシアのトバ火山の噴火による火山灰の冬の寒冷期を迎えると、洞窟内で親密に過ごす時間が増えて、構造主義言語学が「音素」と名づける相互に離散的な一式の音韻単位(論理成分)、クリック子音が生みだされた。音素記憶は新生児に刷り込まれ(刷り込み: 脊椎動物の新生児期や幼少期におきる急速な学習現象で、きわめて安定的で長期的な行動に結びつく)、共同体成員間で共有されて、仲間であることの証明(ID)として機能する。言語は共同体内部の相互認証と意思疎通のために発達した。



図 1 デジタル進化年表

(2) デジタル信号「音節」が生まれる

南アフリカの狩猟採集民が、30 万年前に調理に火を使い始めたことで中期旧石器時代が始まった。13 万年前に彼らは、大西洋とインド洋に面した南部海岸線の分厚い砂岩層

クリック子音は、母音を伴わず、唇や舌を使って出した音を口腔内で反響させる。舌筋を多用するために、下あごの骨が刺激を受けて発達し、下あごが突き出す現生人類に固有の身体的特徴オトガイを得た。

オトガイが下あごの皮膚を突き出させた結果、皮膚と口腔底の間に空隙が生まれて、肺の気道出口が下がっても窒息しなくなった。こうして母音の共鳴を生みだし、水平部分と垂直部分が同じ長さで直角に交わる声道が生まれ、母音アクセントをともなう音節を発するようになった。

音節は、周波数領域で離散的であるうえに、母音波形のおかげで時間軸上でも離散的に処理ができる。ヒトの音声は、音素共有をしていないとひと続きに聞こえるが、音素共有をしていると音節列として聞き取れる。これがデジタル進化の第一段階で、無限の語彙と文法的修飾を生んだ。

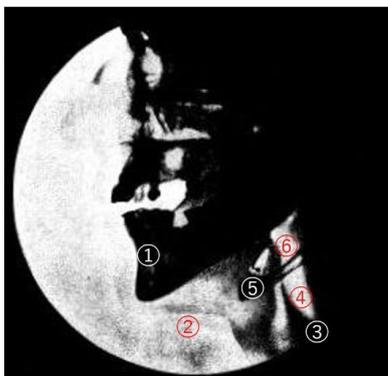


図2 頤の発達により母音共鳴を生む声道を得た[3]
 (① オトガイ, ②皮膚, ③声門, ④喉頭, ⑤舌骨, ⑥咽頭)

(3) 音節・文字・電気信号という3段階の信号進化

文字と bit は、進化した音節である。

文字は今から 5000 年前にメソポタミア地方で発明された。ここはゴンドワナランドがユーラシア大陸に衝突した際に、海として残っていた部分に土砂が堆積して生まれた 1000km x 400km の広大な平原である。この大平原で農耕が始まり、その生産力を基盤として王朝が生まれた。王朝は土地管理と租税徴収のために記録装置を必要とし、文字を発明させた。文字は行政目的で生まれた。

文字とともに学校が生まれ、読み書き能力を教えた。識字者にとって文字列は音声だ。文字列は脳内で内言(inner speech)として響く。文字は「消えない音節」であり、言語情報は時空間を超えて共有され、連続的に発展しはじめた。これが文明である。文字が文明を生み出したのだ。

文明の片隅の僧院で、家族や労働から解放された学僧が、静かな環境で本を読み、徹底的に精思切思した結果、それまで 1 対 1 の反射の論理で意味と結びついていた言葉は、1 対全の群の論理で意味と結びつくようになった。[4][5] 概念については議論が錯綜していて、決定的な定義が存在しないが、言葉記号と意味を結びつける論理が 1 対全の群論理であることがデジタル言語学の「概念」である。

第二次世界大戦が終わるとき、アメリカの科学研究開発局長の V.ブッシュは、MEMEX という研究者のための端末を提唱した。ソ連が 1957 年にアメリカに先駆けてミサイル打上実験に成功すると、アメリカは国防総省に高等研究開発局(ARPA)を設立し、エンゲルバートが指揮するカリフォルニアの研究所でパソコン、インターネット、電子メール、マウス、GUI、コピー、表計算、ワープロ、PDF などの今日の IT 技術の基礎が生まれた。これらすべてがたった一つの研究所で生まれたことは、情報教育のなかで周知させてもよいだろう。[6]

言語学に照らして考えると、電気信号 bit は、処理回路のなかで反応する「対話する音節」であり、検索エンジンにキーワードを入れると、関連する言語情報を瞬時に示してくれる。それまで過去から未来へ一方通行だった流れが、未来から過去へと遡れるようになった。かつて過去の著者が書き残した文書に出会うためには、特殊な文書庫に立入る権利が必要であり、そのうえで徹底的に文書を探して読む集中力と情熱が求められた。今、前半は検索エンジンがやってくれるので、誰でも文書庫に立入れるようになった。

これは人類の文明の営みにとって重大なことだ。有史以来、面々と受け継がれ、発展してきた人類の知的営みが、人類共有知という集合的知識にまとまることを意味する。個別の科学者や哲学者の業績は、それまでの人類共有知を概観して、どれだけの誤りや雑音を正したか、どれだけ真理に近づけたか、分野科学の概念を吟味して学際的に統合したかという点で評価されるだろう。そして有史以来の知的営みは、雑音を除去され、誤りを訂正されて、ひとつの共有知としてまとまり、さらなる発展を遂げることになる。

デジタル言語学は、ヒトの言語処理と知能構築は、脊椎動物の脊髄反射回路を利用して行われる、哺乳類の音声コミュニケーションが基盤にあるとする。それがクリック子音・音節という論理成分(音素)を獲得してデジタル進化が始まり、文字が文明と概念を生み、電気信号の発明によって言語と知能のデジタル進化は最終局面に移った、というきわめてダイナミックな全体仮説である。

表 1 知能の三段階デジタル進化

	引き金	誕生	獲得したもの	場所	時期
1	喉頭降下	音節(音素とモーラ)	語彙数, 無意識の文法処理	南アフリカ	66 千年前
2	農耕と王朝支配	文字(消えない音節)	文明, 科学的概念	大平原	5 千年前
3	総力戦と米ソ ICBM 開発競争	電気信号 bit(対話する音節)	PC, www, 検索エンジン, 前方誤り訂正	アメリカ	20 世紀末

表 1 は、音節・文字・bit という信号の物理的進化がおきたあと、脳内で 1.無意識の文法処理、2.科学的概念(言葉が群の論理で作用する)が生まれたことを示す。

いよいよヒトは、3.前方誤り訂正(過去の著者の雑音や誤りを、受け手が自力で正す)という最終的な論理層進化の時

を迎えた。音素共有と言語獲得の経緯、文明と概念の基盤、脳内回路の生物機構と性質を知ること、正しく学習し、過去の誤りを正す手法を構築して、人類は知的進化の最終段階、知恵ある人（ホモサピエンス）になる時がきたのだ。

2. 前方誤り訂正の人文科学的応用

(1) アナログからデジタルへの進化は雑音を除去する

人文科学においては文献を読み解く教養主義的な研究方法がとられている。しかし文献の真偽の見分け方や、文献に含まれている雑音除去や誤り訂正の手法はまだ確立されていない。デジタル言語学は、そのための手法を提案する。

これまで、私たちは入手した文献にある言葉を、採用するかしないかの all or nothing で二者択一した。アナログ的な受容であり、事実性評価の手法はない。信じるか否かの選択である。

デジタル的には、受け入れる文献に含まれる誤りを回線上で加えられた雑音（通信路誤り、著者の責任に帰すことがない誤り）と、著者が述べた内容の誤り（情報源誤り、著者の責任の誤り）の2種類に分ける。誤りは排中律を構成する2つに区分されるので、通信路誤りと情報源誤りの両方を正せばすべての誤りを正すことになる。

受け取った言語情報がある人物のものであることをまず確認して、続いてその言葉がどのように生まれたかを著者の思考過程を追体験しながら確かめる。文字は消えない音節だから、著者本人の言葉を確認するとは著者の肉声を復元することであり、著者の思考過程を追体験するとは読者が著者になりきって、著者の考察を継承することだ。こうして人類共有知が進歩する。

言語情報の誤りを3種類に切り分ける

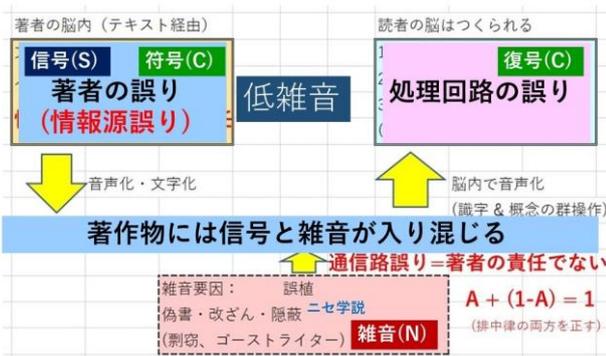


図3 言語情報には信号と雑音が入り混じる

(2) 通信路誤り訂正：書誌学と歴史的なアプローチ

テキストから著者の肉声を聞くためには、書誌学の知識が必要となる。著者本人が望む通りの文字列を復元するための技術である。

テキストの真実性は、著者の生きた行跡が保証する。誰が書いたのか、その人はどんな人物で、どのような人生を送

ったのかということが重要である。これは歴史学である。

(a) 著者存命中に出版された書籍や論文は信号(S)

著者存命中に出版された書籍や論文は、著者本人が著者校正に携わっている可能性が高く、肉声とみなしてよい。したがって本は、著者の生没年と、初版の出版・改訂時期を見比べる必要がある。

プレート理論を独立した一つの章として取り込んだ嚆矢として知られるアーサー・ホームズの「一般地質学 原書第三版」は、アーサーが第二版を改訂してまもない1965年に没した後、1978年に妻であったドリス・ホームズが共著者に加わって第三版が出版されたものである。[7]

著作権者であるドリスは「内容に何か欠点や誤りがあれば、それはすべて著者の責任である」と宣言しているが、彼女は初版や第二版では共著者ではなかった。第二版から第三版の改訂箇所の特記も趣旨説明もない。しかしこの第三版によって、科学的な評価や検討のないプレート理論が世界中で受け入れられるようになったのである。今さらだが、一度原点に戻って考え直す必要があるだろう。

(b) メモや未公開の手稿からも声は聞こえる(S)

メモ書きや印刷に至っていない手稿でも、日付や筆跡などによって本人のものであることを識別できれば、本人の考え方の遷移を確認できる。

筆者は米国議会図書館の手稿閲覧室で、1949年にフォン・ノイマンがイリノイ大学で行った連続講義のタイプ打ち原稿を閲覧した。イリノイ大学がタイプ打ちしたものを送ってきていたもので、出版のために送り返すよう大学出版局から督促状も来ていた。[8]

講演(手稿)のなかでフォン・ノイマンは、情報理論を説明するにあたってマックスウェル流のエントロピーを持ち出している。彼は1949年の時点で情報理論エントロピーをマックスウェル流に説明していたが、その後の出版物のなかでは熱力学的で説明している。これは彼が考えを改めたことを示唆する。

(c) 剽窃とゴーストライター作品の排除(雑音N)

ある本や論文の前後に類似のテーマで講演や論文があるかどうか、著作の内容について著者が誠実に質問に答えているかどうかは、剽窃やゴーストライターの可能性を検討するうえで重要である。

剽窃やゴーストライター作品の問題は、誰もそのテキストに責任を持たないところにある。真正の著者は自分の名前が残る言葉に、できるだけ誤りがないように細心の注意を払う。フォン・ノイマンが(b)のイリノイ大学講演録を出版させず、情報理論のエントロピーは熱力学概念であることを示唆した言葉を遺したのは、自分自身のプライドもあったろうが、むしろ人類共有知への貢献という責任感に由

来したと思われる。

剽窃は一時的に偽の業績として使われるだけで、発展性がない。ゴーストライター作品は真の作者も依頼主も内容の真実性を追い求めない。いくら立派なことが書かれていても、著者の責任に裏付けられていない言葉は使い物にならない。使ってはいけない。

クロード・シャノンは、1938年3月に提出したとされる修士論文のなかでリレースイッチ回路を使った解析を論じているが、その前後に類似の論文がみられない。[9]後にアメリカの科学雑誌OMNIが行ったインタビューのなかで、ON/OFFのリレー回路を使ってYES/NOの論理表現を行うことをどうやって思いついたのか、それはひらめきではなかったかと何回か聞かれたが、ひらめきではないと繰り返し、おしまいには「ひらめきとは何であるのか知りません」と言っている。シャノンという人物への疑問が募る。[10]

一方で、日本の中嶋章は3年前から同じ継電器回路の問題について1935年から37年末にかけて8本、1938年から41年にかけて5本の講演と論文を電信電話学会誌上で発表している。[11][12]

中嶋は1970年に「1939年から40年にかけて米国を出張し、A. I. E. E. の Winter Convention に出席したところ」、「M. I. T. の C. E. Shannon 氏に紹介され Switching Circuit Theory について話し合ったが、そのときの同氏の若々しい理知的な顔立ちはいまだに忘れられない。」と報告している。[13]そしてシャノンと話し合った後(1941年)に「近年独逸及び米国に於て一二の文献が発表されたが其の理論的根拠は筆者等の方法の踏襲に過ぎない。」と書いている。[14]

現代の電子デジタル技術のもっとも根源にあるリレースイッチ回路について、公開情報でも多くのことがわかる。曖昧さや矛盾をひとつひとつ丁寧に解析し解明することで、さらなる発展が可能になるのではないか。

(d) 真理ロンダリング：裏を取らない無闇な引用(N)

一般に学者や研究者は、自分の専門分野に属する研究実績や著名論文からの引用については目を光らせて妥当性を判断するが、大衆向けの科学書を書くにあたって、専門外の研究書や論文から引用する場合、その研究が妥当であるか、どのような実験や観察にもとづいてどのように判断したのかというあたりをまったく調べないで結果だけ紹介することが多い。これは「科学的真理ロンダリング」であり、雑音レベルを高め、誤解や誤りを広める結果を導いている。

新聞やテレビなどの大手マスコミが、Cell や Nature や Science に掲載された記事を吟味せずに紹介するのも、同じく誤解や誤りを広める効果をもつ。真理ロンダリングは大衆の誤解を広める効果をもつため、人類の文明発展という見地にたてば、慎むべきことである。啓蒙的科学書のいかがわしさを、きちんと批判するためには大変な手間と時間がかかる。

自分自身が研究者なら、安易にマスコミや啓蒙科学書の真理ロンダリングにひっかからないように注意する必要がある。筆者は2014年春にSTAP細胞がマスコミで注目された(炎上した?)とき、大学図書館で雑誌Natureを複写して、小保方晴子さんの2本の記事を読んだ。gfpが何の略語か説明なしに使われていることにまず驚いた。

小保方さんと笹井教授たちの実験は、マウスのBリンパ球をオレンジジュースなどの刺激物に浸けてから、別のマウスの子宮に入れると、3日ほど活性化するというものだった。これは、利根川進博士が1987年に行ったノーベル講演で論じた、免疫二次応答でおきる体細胞超変異(somatic hypermutation)と似ている。彼らとそれを議論したかったが残念ながらできなかった。[15]

『有無の無はしばらくおく、いかならんかこれ仏性』と問取すべし、『なにものかこれ仏性』とたづぬべし。(正法眼蔵仏性)と道元は教える。[16]「仏性」を「STAP細胞」に置き換える必要がある。それが何かを確かめずに有無を論ずるのは有史以来の悪癖であり、克服する必要がある。

真理を探究するために、本や論文をきちんと読む必要があることは、誰もが認める。2020年6月の回路とシステム研究会で、「そんなことをすると本を読むだけでも30年かかる(だからできない?)」というコメントを頂いた。[17]このコメントは悲観主義的すぎる。一人ですべてやろうとすると時間はたしかに足りないかもしれないが、みんなで協力すれば、すぐに雑音や誤りの多い本や論文は淘汰されて、科学はさらに発展するだろう。

(e) 翻訳の妥当性判断は難しい(N)

翻訳はある言語の音韻刺激を別の言語の音韻刺激に変換するものであり、脳外の物理層の現象である。そこで生まれる誤りは著者の責任ではないので、通信路誤りに属す。

ある言語から別の言語に翻訳される時、両方の言語に通じていて、双方で言葉が想起させる意味を理解していないと、訳語の適切さは判断できない。そのため長い年月にわたって間違った訳語が通用することがある。

たとえばチョムスキーはフランス語のlangageを英語のlanguageに翻訳する。見た目も似ていて、問題ないようにもみえるが、彼はフランス語のlangueもlanguageに訳している。languageに対応するのはむしろlangue(言語)なので、langage(言語表現、言葉)をlanguageに訳すのは誤解を生む。

langueを言語、langageを言語表現として別々の概念として用いているソシユール言語学は、チョムスキーの訳語を使うと、二つの異なる概念が英語では一つの概念になるために大きな混乱を生じる。なぜこのような翻訳慣行が50年間許されてきたのか疑問である。国際的な対話不足のためだろうか。

一般読者が訳語選択の妥当性を判断するのは容易では

ない。矛盾や誤りに気がついた人は、そのことを公表して、広く周知させる必要がある。そして、その報告を受けたら、学会や研究者は矛盾や誤りに真摯に取り組んで、ひとつひとつ正していくというルールを作る必要がある。

(f) 没後の改ざんや偽書から真筆を保護する符号(C)化

日本の中世には偽書が多いが、「世に流布する文書の真偽を厳密に区別しようとする指向性は皆無」であり、「ひとたび作り上げられた偽書は、聖徳太子未来記にみられるように、驚くほどすんなりと社会に受け入れられ」た。そのため「膨大な数の偽書が流通」している。[18]

これを正そうという動きはみられない。正す方法はない、どうにもできないと思っているからだろう。死人に口なしという諺どおりに、偽書が流通しても死者はそれを指摘できない。なんとなく怪しい文書であると感じても、後世の人間がそれが偽書であることを証明することは難しい。

しかしデジタル通信の誤り訂正符号の理論にもとづいて、著者が自身の著作に冗長的な符号を付加していたなら、符号と真筆の整合性によった保護が可能である。著者がある法則にもとづいて符号化を行って符号を付加した場合、読者は符号化の事実と法則を理解すれば、符号にもとづいて正筆を識別できる。

デジタル通信の誤り訂正符号は 20 世紀後半に考案された。ハミング符号やギャラハの LDPC (Low Density Parity Check Code, 5G で採用)、ヴィダヴィ符号などが有名である。しかし世界で最初に自著を符号化して保護を試みたのは筆者が調べたかぎり、13 世紀の日本の学僧道元 (1200-1253) である。[19]

道元は、(i) 不勉強な日本達磨宗の弟子たちが自分の教えを封印・改ざんし、偽書をつくると予見して、自著を符号化して保護する必要があった。

そして(ii) 『道元和尚廣録』の上堂語 473 で「嵩岳(すうがく)の高祖云へり「我が滅後八千年、我が法は糸髪(しはつ)の如く許りも移らず、我が在る世の如くならん」(達磨大師は自分が死んだ後 8000 年、自分の説いた法は一切変わらないと言った)と語り、(略)「臘月の寒梅、月光を含む。雪山の雪の上に、更に霜を加ふ。如来の毫相、猶、今に在り。遠孫を利益す、豈に度量せんや」という頌(漢詩)を残している。「雪の上に霜を加える」は、冗長符号を添付したことの示唆であろう。[16]

(iii) 『正法眼蔵』の 75 巻本(と弁道話)は連番管理されていて、それぞれ奥書のなかでいつどこで講義し清書したかが示されている。例えば弁道話は「寛喜辛卯中秋日」(1231 年 8 月 15 日)、正法眼蔵出家第 75 には「正法眼蔵第七十五、爾時寛元四年丙午九月十五日在越宇永平寺示衆」(1246 年)であり、残りの 74 編はこの 2 時期の間に示衆されている。

干支だけで書かれていて年号のないものもあるが、符号化の事実や意味に気づかせないためであろう。正法眼蔵現

成公案第一の奥書に「建長壬子拾勒」とあるのは、道元が亡くなる前年の建長 4 年(1252)に正法眼蔵の第一巻は収録されたのであり、道元が亡くなる直前まで執筆をし、残りの巻もすべて手を入れたことを示す。筆者が偽書と断定する新草 12 巻本の巻 3 奥書が書写日とする建治元年(1275)の干支(乙亥)を翌年の丙子に間違えている。これは自ら偽書であると宣言したことになる。

こうして正法眼蔵 75 巻本の連番と示衆日・書写日は結界を築いて、不用意に偽書を混入させたり削除できない構造をつくっている。

一方、『道元和尚廣録』は十巻あるなかで 531 の上堂語を収録する第 1 巻から第 7 巻までは、各巻の上堂語数と頌の数合計した数字が識語として書き込まれている。漢文白文で書かれた本文から、漢詩を数え上げるのは著者以外には不可能だ。おかげで上堂語の削除や、偽の上堂語を挿入しにくくなっている。複数の写本が存在するとき、読者は識語をもつ方が真筆である可能性が高いとわかる。

(iv) 真筆は保護されているが弟子の裏切りは起きた

道元は真筆を符号によって保護したが、弟子は 75 巻本を否定するいわゆる「新草 12 巻」をつくり出した。この 12 巻は、どれも示衆日がなく、書写日も「結界」の外にある。道元らしからぬ文体と内容であるにもかかわらず、曹洞宗と学会は新草 12 巻が道元の作であるという扱いを続けてきた。書誌学と誤り訂正符号の知識があれば、簡単に見破れる偽書であるのに残念である。

道元の主著である『正法眼蔵 75 巻』と語録の『道元和尚廣録』はこれまでほとんど読まれてきておらず、代わりに似ても似つかぬ内容の『正法眼蔵随聞記』が道元の語録として流布している。道元が符号化によって真筆を保護したことを理解するために、道元研究者は誤り訂正符号化の勉強をする必要がある。

入手した言語情報(テキスト)が、著者の真筆でないことがわかったら、読書対象から除外する。完全に除外できなければ、優先度を下げ、参考資料程度の扱いにする。

(3) 情報源誤り訂正：著者になりきって再生的に知識獲得

(a) 文字情報を仮想現実として読む

通信路誤り訂正は、冗長性を利用して、書誌学的・歴史的に真筆を識別する。我々は、こうして著者本人の声を聞くことになる。

情報源誤りの訂正はより難しい。なぜならば、読者よりも著者のほうが物知りであり、その分野での研究歴も長い。著者の知識と同程度の専門知識をもたなければ、著者の伝える情報は消化吸収できない。

先入観を排することは大切である。しかし読み解くうえで必要な知識はもっておかなくてはならない。そして著者の言葉を無心に受け止めて、実験でおきたこと、著者の観察、著者の思考を、自分自身の観察・実験・思考であるか

のように仮想現実を再構成することがのぞましい。

(b) 本や論文はまず3回読む、新分野はレビューから

いったいどうすれば、学際的に本や論文を読んで、正しいかどうかを判断し、誤りを認識してそれを正し、さらには発展させることができるのだろうか。

自分の知らない分野の本は、たとえそれが良い本であっても(良い本であるほど)なかなか入門者の頭の中に入っていない。自分にとって新しい分野の成果を吸収するためには、一定期間チンプンカンプンな言葉と取っ組み合いする必要がある。脳内に言葉の記憶ができて、思考を重ねるとチンプンカンプンな状態は改善する。

レビュー(発展史や課題を示す概観論文)論文は、その分野の最近の研究動向をまとめている。これを読んで、方向性を理解し、重要単語を覚える必要がある。自分が知らない単語に出会ったら、きちんと辞書を引いてその単語の読みと意味を理解し、自分の脳内にその単語の受容体を形成する。言葉なくして思考なし。言葉の記憶を構築してそれを自由に操れないと、思考はできない。

本や論文を読む時は、少し時間を空けて三回読むのがよい。一回目は著者が何をどう語っているか、二回目はそれに対して自分はどこまで対応できるかと記憶を振り返る、三回目は著者と自分が対話する。こうすると、良い本か悪い本かもみえてくる。悪い本との付き合いはここでやめる。

良い本は、研究会や学会で取り上げるなどして、できるだけ繰り返し読む。読むスピードや精度や範囲を変えて、読む。著者の気持ちになりきる。

(c) 著者が解明し残した謎をみつける

基本的に読書は著者が解明したことを学び、継承するためである。しかし時々著者が解明できなかった謎や難題を継承することもある。

誠実な著者であっても、プライドが高く、解明できなかった謎をそのまま書き残さない。したがって読者は一度や二度読んだくらいではその謎に気づかず、何度も読み込んだときによくそれがわかる仕組みになっている。

例えば道元は正法眼蔵光明第15の奥書で、「大家、未だ雲門の道を戯破するを免れず」(諸君、まだ雲門の言ったことを見破る仕事が残っている)と記す。道元は「僧堂・仏殿・厨庫・山門」という雲門の言葉が理解できなかった。それは後世の君たちの仕事だと言い遺しているのだ。[16]

パブロフは、条件反射実験を報告する連続講義の第19講から第21講の3回をイヌの脳皮質の切除実験にあてている。読むだけでも痛々しくて、はじめのうちは軽く読み飛ばしていた。パブロフが多数のイヌの命を犠牲にした意図も理解できないでいた。[20]

よく読むとパブロフは第2講で「大脳半球のない犬」は、刺激に対して決して唾液を出さないだろう(上 P46)と述べ

ている。これを確かめる実験だったのだ。ところが「手術が大脳半球のどの点であろうと、手術後普通無差別に条件反射はみな消失する。消失期間はさまざまで一日から数週間、時には数カ月に及んだ。」(下 P124)と第19講に書いている。つまり条件反射は大脳皮質を切除しても回復した。仮説と結果が矛盾するため、両者を離れた場所におき、誠実にも「私が報告したデータの中にもかなりの失敗があると思っている。」(下 P186)と告白している。

デジタル言語学は、言語は大脳皮質ではなく、脳室内免疫細胞ネットワークが処理しているという仮説である。この仮説には、パブロフの研究成果が生かされている。[21]

(d) 著者の知識獲得プロセスを再生的に受容

デジタルシステムは、「自動的に複雑なものを再生し、時には進化を生み出す」ところに真骨頂がある。たとえば、哺乳類の生殖は、ひとつの受精卵が細胞分裂と複雑化のプロセスを繰り返して、独立したひとつの生命体に成熟する。

「個体発生は系統発生を繰り返す」とするヘッケルの法則によって、原始生物から魚類、両生類、爬虫類の時代を経て、哺乳類になる。初期人類進化である直立二足歩行や現生人類のオトガイの発達はかなり後ろのほうで(一部は母胎を離れた誕生後に)発現する。

おそらく進化の道筋を歴史的にたどり、新たな進化がおきるたびに枝分かれしていく「再生的」発生が、自動的に複雑なものを再生し、時には進化を生み出すのに最適であるのだろう。

知能の進化はどのようにして生まれるのだろうか。新たな物質や現象の発見とその命名が、知能の進化をもたらすといってもよいのではないだろうか。そしてそれは集団的に生まれるのではなく、ある個人が発見・開拓するものだ。

たとえば地動説は、コペルニクス説と呼ばれている。それまでは太陽が地球の周りを回っていると考えられていたところを、コペルニクスが反対に地球が太陽の周りを回っていることに気づいたのだ。また、遺伝学はメンデル理論と呼ばれるように、メンデルが行った実験から発展した。

すると知識再生型の受容を目指すためには、その現象の発見者や概念の命名者の本や手記を読み、いつどこでどうやってそれを発見したのかを追体験することがのぞましい。

過去に発見された科学的知識を自らの記憶に加えることは、さらなる科学的発展を産み出すことにつながる。それは自分自身を、人類共有知を産み出す集合体としてのホモサピエンスの一員に高める行為である。

(e) 著者が概念を産み出した瞬間を見極める

インターネット検索は、きわめて大量だが信頼性の保証のない学際的な言語情報を探し出す。①大量、②信頼性、③学際という3つの課題を乗り越えて情報を使いこなすための手法(読書法)は、まだ手探りの状態であり、構築され

ていない。信頼性という点では、(i) 著者自身で考えた言葉か、(ii) 誰かから聞いた(受け入れた)言葉かを識別し、(i) ならば著者の本を丁寧に読み、著者がその概念を発見した経緯と発見の瞬間を追体験する。(ii) ならば、参考文献にあるオリジナル発見者の本にそれを求める。

こうすれば、これまで誰も説明できず、鵜呑みにするほかなかった複雑な科学的概念が、具体的な概念として理解できるようになる。また、熱力学と情報理論で同じ「エントロピー」という概念の意味が異なるとされてきたが、それぞれ概念の起源を求めて遡ってみれば、そんなことはありえないことがわかる。

筆者はエントロピーという概念の定義を求めてマックスウェル、ボルツマン、プランクを読んだがみつからなかった。それを産み出したのはポーランドの物理学者クラウジウスであり、彼の力学的熱理論の全体像を理解しようとする八木江里氏が、クラウジウスが最初の論文集に収録した 17 論文から全ての方程式(約 500 個)を取り出して分析し、「物体の分散度」という新しい量を導いた。これがエントロピーであり「構成要素の配置の変化量」である。これは、変化量、歪み量だから正の値しかもたず ΔS という表記は妥当である。[22]

一方、情報理論では、誰がどのようにしてエントロピー概念を産み出したのか明らかでない。どこかで伝言ゲームとなり、意味が逆転してしまったとしか思えない。

科学的概念と出会った人は、自分がこれまで思いもしなかった世界を発見して驚く。そして発見の後、その概念の発見者の足跡をたどり、多くの後進研究者たちによる発展をおいかける。ひとつの謎が解明されると、新たな謎がみえてくる。科学的概念の意味にこれで終わりという領域はなく、常に前進し新たな関係性を生みだしている。したがって常に自分自身を前衛として意識して、自分に課された仕事を理解しておかなければ、科学的概念は正しく受容できない。自分の研究分野でないことであっても、自分の保有している知識は、現在わかっていることのすべてであるかを確認する。すべてを知る必要はないが、すべてではないことを知っていれば、必要に応じて、不足した部分を探し求める身体をもつことができる。

(f) 言葉の意味に 6 種類ある複雑化のメカニズム

核内 DNA の二重螺旋構造に記された情報だけにもとづいて、単細胞の受精卵が独立した生命体へと複雑化するメカニズムを、生物学はまだ解明していない。生物学の分類法の伝統はそれを求めてすらいけない可能性がある

一次元情報が自立した三次元の生命体になる複雑化メカニズムを解明するのは、情報理論の役目である。フォン・ノイマンは「人工頭脳と自己増殖」の第 7 章「複雑化の概念：自己増殖」のなかで、複雑度を増加させる循環論法のメカニズムに言及している。これはフラクタルではないか。

複雑次数をひとつ上げる(ひとつ複雑度を増す)式が存在し、その式のおかげで複雑化したものを、同じ式に投入することでもう一段階複雑になる。

言葉と記憶を結びつける論理が、1 対 1 の反射の論理から、1 対全の群の論理となるとときに複雑化がおきて概念が生まれる。この二元的な構造に、概念を投入すると、全対全の複雑なネットワークの論理へと進化する。言葉の意味の複雑化のプロセスはこれまで解明されてこなかったが、フラクタルな構造に求められるのではないだろうか。

3種の論理と2種の記憶が意味を構築

言葉と記憶を結びつける論理	(a) 五官記憶	(b) 思考操作の記憶
(i) 1対1 反射	特定の五官記憶の想起	科学的記号: 恣意的命名 (科学的な現象)
(ii) 1対全 群	生活概念: 五官記憶の総合化	科学的概念: 学際的統合による科学的現象の総合化
(iii) 全対全 複雑なネットワーク	類と関係性の概念 (生活概念の概念操作による)	学際科学概念 複数の下位システム相互の自動的なネットワークワーキング

表 2 意味の複雑化を生み出すフラクタルな構造

(g) 概念が群であることを確認する

概念が群であるときに、群と群の演算だから意味ある結果を生み出す。概念を正しく獲得し使うために、ピアジェは 5 つの条件(I)~(V)を示している。詳細は MPS 研究会で報告したのでそちらを読んでいただきたい。[5]

- (I) 合成性 $x + x^1 = y; y + y^1 = z; \text{etc.}$
- (II) 可逆性 $y - x = x^1 \text{ or } y - x^1 = x.$
- (III) 連合性 $(x + x^1) + y^1 = x + (x^1 + y^1) = (z).$
- (IV) 同一性 $x - x = 0; y - y = 0; \text{etc.}$
- (V) 同質性 $x + x = x; y + y = y; \text{etc.}$

新しい概念に出会うたびに、この条件に照らして群の条件を満足しているかを確認すると、学際的研究は混乱することなく前進するであろう。

3. おわりに：人類とは共有知である

南アフリカの洞窟の中で、音声の離散成分(音素とアクセント)を獲得して、ヒトのデジタル進化は始まった。それが文字と bit の発明によって、発明者の予想を超えたダイナミックな展開を遂げている。この三段階進化の複雑度の増え方もまたフラクタルである。

文字による知能進化第二段階の成果である文明は、文字化された言語情報が時空間を超えて共有され、連続的發展を遂げる現象である。概念は、1 対 1 の反射の論理が 1 対全の数学的な群論理へと進化したものである。これら文明と概念の本質を理解して、人類共有知の発展の系譜に自らを位置づけているのは、ごく一握りの人々であった。

Bit とコンピュータ・ネットワークによって始まった第三進化は、人類共有知のデジタル進化を完成することになる

であろう。言語情報に含まれている誤りや歪を伝えるジャンク遺伝子が淘汰され或いは不活性化されて、共有知をまっすぐ発展に導く良いゲノムだけが継承されることになる。

ヒトが言語処理と知能構築に、不随意メカニズムである脊椎動物の脊髄反射回路を用いていることと、音素共有した背景に共同体として生きる真社会性動物であることが作用して、知能進化には雑音や歪が混じり、大いに混乱しているのが、有史以来の科学と哲学である。個人の無責任や共同体のゼノフォビクな性質で広められた嘘や偽りは、通信路誤り訂正と情報源誤り訂正によってすべて暴露され消えていくことになるだろう。

現生人類とは、哺乳類のアナログな鳴き声記号が、周波数領域と時間軸上の離散成分を含む音声へと進化したことによって、無限の語彙、文法、時空を超えた共有と連続的発展、科学的概念、キーワード検索、言語情報の前方誤り訂正という複雑な知能の進化の道筋に達したサルである。言語獲得によって現生人類は誕生したのであり、三段階のデジタル進化を達成することで人間になれる。言葉を正しく学び、正しく使うことによって、ヒトは人間になれる。

知能の生物メカニズムの不随意的な制約と、音素共有の真社会性バイアス（ゼノフォビアや上意下達の悪しき伝統）を乗り越えることで、ヒトはようやく知恵ある人（ホモサピエンス）になるのであり、嘘や偽りや誤魔化しは、人間になることを放棄することにつながるということを、すべての人類が理解する必要がある。個人の目指すべきものは、人類共有知の発展にどれだけ貢献したかということになり、そのために寸暇を惜しんで勉強する人間が増えることが、人類の理想的な未来像である。人類とは共有知であり、その発展に個々人の存在理由は求められる。

その際に、人文科学の書誌学や歴史学は真筆を見極めるうえで、また著者の人となりを知り、概念誕生の瞬間を見極めるうえで、読み書き能力と同じくらい重要な知識として、人間の身体の一部を構成することになるだろう。人文科学のさらなる発展に期待する。

謝辞 発表の場を与えてくださった人文科学とコンピュータ研究会の皆様へ感謝申し上げます。

参考文献

- [1] ポルトマン, A.(1961) 人間はどこまで動物か 新しい人間像のために, 高木正孝訳, 岩波書店
- [2] ウィルソン, E.O. 人類はどこから来て, どこへ行くのか, 斉藤隆央訳, 化学同人, 2013 年
- [3] 千葉勉・梶山正人, 母音—その性質と構造, 岩波書店 2003
- [4] 得丸 群論理に基づく概念操作-言葉記号から概念への進化 (デジタル言語学) 信学技報 HCS2019-54
- [5] 得丸 概念に求められる群性 概念の意味の前方誤り訂正のために IPSJ 2020-MPS-130 No.2
- [6] 得丸 デジタル言語学—知能発展の第三段階 E-017 第 19 回情報科学技術フォーラム (FIT2020)
- [7] 得丸 月誕生のカラハリ隕石衝突仮説の発展 - 19 世紀地質学

- の継承, 写真測量とリモートセンシング 54:216-220
- [8] 得丸 情報理論のエントロピー概念の再構成- シャノンの伝記とノイマンの手稿を参考に, 信学技報 RECONF2019-35
 - [9] Shannon C.E., A Symbolic Analysis of Relay and Switching Circuits, Transactions American Institute of Electrical Engineers, Vol. 57 (1938), pp. 713-723. (Received March 1, 1938.) Included in Part B. Claude Shannon Collected Papers, IEEE 1993 PP471-495
 - [10] Riversidge A., Profile of Claude Shannon, Claude Shannon Collected Papers, IEEE 1993 xix-xxxiii
 - [11] 得丸 自然とひらめきが生まれる学習法, IPSJ-CH14102001
 - [12] 中嶋章 継電器回路の構成理論, 講演, 電信電話学会雑誌第 150 号 (1935 年 9 月)
 - [13] 中嶋章 スイッチング回路網理論の思い出, フロンティアを求めて, 電子通信学会誌 (1970 年 12 月)
 - [14] 中嶋章 継電器回路の理論, 通信技術展望, 電信電話学会雑誌第 216 号 (1941 年 3 月)
 - [15] 得丸 科学論文を読むためのセンシング技術～ 真理を伝える言葉を感じ取る ～ 信学技報 SeMI 2020-15
 - [16] 得丸 道元を読み解く, 富山房インターナショナル 2017
 - [17] 得丸 デジタル言語第三進化 (電子情報) 論理層の構築～ホモ・サピエンスになるために 信学技報 CAS2020-1
 - [18] 佐藤弘夫 偽書の精神史, 2002 講談社
 - [19] 得丸 道元(1200-1253)が正法眼蔵と道元和尚廣録に施した誤り訂正符号～原著者の真正なテキストが 21 世紀に蘇る 信学技報 COMP2018-8
 - [20] パプロフ 大脳半球の働きについて, 岩波文庫 1975
 - [21] 得丸 言葉認識と意味理解の生物構造: 言語は大脳皮質で処理されていない (デジタル言語学) 人工知能学会 SIG-SLUD-089-10
 - [22] 得丸 タンパク質の情報理論 — 発癌メカニズムのエントロピー仮説 FIT-2015 G-005 (第 14 回情報科学技術フォーラム)



図 4/5 クラシーズ河口洞窟 3 号 (入口・内部)



© Tokumaru Kumon 2007/2012