

デジタル言語学 — 知能発展の第三段階 Digital Linguistics – The 3rd Stage of Intelligence Development

得丸 久文[†]
Kumon Tokumaru

1. はじめに: 言語的人類の起源を求めて

筆者は 2007 年に、南アフリカ共和国にある最古の現生人類洞窟、クラシーズ河口洞窟を訪問した。南アフリカ政府のアパルトヘイト(人種隔離)政策を口実として、1960 年から 1990 年代まで徹底的な対南ア・国際学術ボイコットが行われていたため、この洞窟をはじめとして南アの人類学と考古学遺跡に関する研究はおどろくほど少ないし、研究者数もきわめて少ない。初期人類アウストラロピテクスアフリカヌスの化石も、現生人類化石も、南アフリカで発見されたことから考えると、異常である。とりあえず自分の目で現生人類が生まれた場所を確かめてみよう、私費で一人旅を企てたのだ。

クラシーズ河口洞窟の第 3 号洞窟は、インド洋と大西洋が交わる波の荒い海域に面した、砂岩の海岸段丘にある。海拔 20m に穿たれた海蝕洞窟は、広くて天井も高く暖かく、快適な居住環境だった。もしここで人類が言語を獲得したのだとすると、洞窟内では夜は真っ暗になるので音声コミュニケーションが盛んになったためではないかと、洞窟進化仮説(2008 年日本進化学会で発表)をたてた。[1]

ところが、もし暗闇が言語進化を引き起こしたのなら、地下トンネルで生活するハダカデバネズミ(真社会性で、晩成性で、皮膚も毛も薄い)の音声記号数も多いはずだと思っ専門書を調べたところ、記号の使用頻度は多いが、記号数は地上生活のデバネズミと同じだった。[2] 洞窟と語彙数は相関しないことがわかり、洞窟進化だけでは説明がつかないことがわかった。

そのとき、ひらがな積み木セットが脳裏に思い浮かび、音節の順列のおかげで無限の造語力が得られることに思い至った。2 音節なら音節数の 2 乗、3 音節なら 3 乗と、桁(digit)のおかげで順列できるのはデジタル(digital)かと考えて、2008 年秋に情報処理学会に入会し論文にまとめて提出したが、2 度連続して D 判定(不採択)になった。そこで 2009 年秋から電子情報通信学会と情報処理学会の研究会に 100 回以上参加させていただき、言語進化について学際的な検討と統合を続けてきた。

	開催日	鈴木孝夫研究会のテキスト [3] (表 1)
第 1 回	2010.3.20	記念講演「青山墓地を世界遺産に」
第 2 回	2010.5.29	『ことばと文化』
第 3 回	2010.7.10	『私の言語学』
第 4 回	2010.9.18	『教養としての言語学』
第 5 回	2010.11.20	『日本人はなぜ英語ができないか』
第 6 回	2011.2.12	『新・武器としてのことば』
第 7 回	2011.4.29	『日本語と外国語』
第 8 回	2011.7.9	『言葉のちから』
第 9 回	2011.9.17	『閉ざされた言語・日本語の世界』
第 10 回	2011.11.27	『人にはどれだけの物が必要か』
第 11 回	2012.2.4	『日本語と外国語』(第 4 章, 5 章)
第 12 回	2012.4.14	『日本人はなぜ日本を愛せないのか』

[†] 著述業 Writer (e-mail) tokumaru @ pp.iij4u.or.jp

同じ頃に言語学者鈴木孝夫先生とのご縁が生まれ、研究会に参加し、意味論、空の記号、分節、動物生態学など、鈴木先生の著書や論文から多大なる教えを受けた。[3]

2. 研究手法: 古典精読と進化の起きた現場を歩く

2.1 検索エンジン⇒OPAC⇒精読という流れ

2008 年秋に Google に“human, digital, language”と入力して「ヒト言語のデジタル起源」という論文を知った。[4]この論文には 56 の参考文献があった。シャノン、ピアジェ、クリック、ポルトマン、チョムスキー、ピンカー、キャヴァリスフォルザ、ノイマン、メイナードスミス、イェルネと、情報理論・分子生物学・言語学・免疫学・進化生物学の碩学と古典を知り、それらに目を通した。

ピアジェ、フォン・ノイマン、イェルネは明解な語り口だが内容は理解しづらかったため、何度も研究会で取り上げ、繰り返し読んだ。できるだけ静かな環境で、信頼性の高い言語情報だけをくり返し精読した。

一方、チョムスキーやシャノンの著作に、あいまいさを発見し、それを掘り下げることが新たな発見を生んだ。

2.2 脳室内免疫細胞ネットワークによる言語処理仮説

言語処理メカニズムを明らかにするために、筆者もはじめは通説にしたがって大脳皮質で言語処理されているという前提で考察を続けていたが、おどろくべきことにそれと合わない実験結果をペンフィールドやバプロフが書き残していることに気づいた。[5][6] さらにイェルネの免疫細胞ネットワーク理論から、脳室内を浮遊する B リンパ球は神経細胞がモバイルネットワークするように進化したモバイルニューロンであり、言葉記号の処理に必要なすべての要素をもっていることを知り、「脳室内免疫細胞ネットワークによる言語処理仮説」が 2012 年に生まれた。[7]

2.3 文字と電子化が人類の知能にどう影響するか

イギリスの生物学者ジョン・メイナード・スミス(1920-2004)も言語のデジタル性を指摘していた。彼は、音節を獲得した後、文字が発明され、情報の貯蔵と伝達のために電子的手段を使う進化がおきたと指摘する。[9]

筆者は 2016 年にパキスタン国際言語学会に参加し、古代文明の生まれたインダス川平原を訪れた。そして首都イスラマバードから学会が開かれたパンジャブ大学のあるラホールまで約 300km、丘ひとつなく水利の良い農地が延々と続いていることにおどろいた。帰国時のフライトマップで、インダス河平原は、ゴンドワナ大陸の一部であるインド亜大陸がユーラシアに衝突した際に海として残っていた部分に、ヒマラヤ山系からの土砂が流入して生まれた土地であることを知った。

ヒトの認知能力を超える大平原で文字が発明されたのは、王朝が徴税のための土地の記録を作成するためだったのだろう。四大古代文明はすべて大陸間大陸亀裂で生まれている。文明が文字を生んだのではない、広大無辺な地形が記録装置を必要として、王朝が文字を発明させ、それを使って言語情報が時空を超えて共有されるようになり、世代間での知識の継承と発展が生まれた。文字が文明を生み出したのだ。

文字と文明の関係が逆転したとき、音節と文字の関係も違ってみえてきた。文字列は、識字者の脳にとっては音節列(音声)である。文字のおかげで音節は消えなくなり、時空を超えるようになったのだ。

では「情報の貯蔵と伝達のために電子的手段を使う進化」は何を意味するか。電子検索による双方向通信の実現ではないだろうか。文字だけの時代、過去に書かれた書物や論文に出会うためには、図書館で雑誌のバックナンバーに片っ端から目を通し、蔵書カードを一枚一枚めくって出会う必要があった。あとは書店や古書店の店先で渉猟する方法であり、運や縁や根気を必要とした。

電子化によるキーワード検索のおかげで、テーマや著名名さえわかれば、文献とその所在地が容易にわかるようになった。おかげで「**きわめて大量だが信頼性の保証のない言語情報**」を入手できるようになった。

①文字の獲得がどのように人類の知能を発展させたのか。
②電子化のおかげで得られる言語情報はどう使えばよいのか。本稿はデジタル言語と知能の物理層と論理層での処理手法の三段階の進化をみきわめることによって、この二つの問題を明らかにすることを志すものである。

3. 第一進化：音節の獲得と文法処理

3.1 頤の発達で喉頭降下に結びついて音節を獲得

哺乳類は今から 6600 万年前に誕生し、鳴き声の波形を記号とした音声コミュニケーションをはじめた。ヒトの言語は、哺乳類の鳴き声の波形にもとづくアナログ記号が、音素(phoneme)の順列によってつくられるデジタル記号に進化したものだが、それは出発点であり、音節は文字や電子化へと進化し、脳の使い方もより複雑に変わっていく。

また、今から 7 万 2000 年前に始まるスティールベイ文化は、トバ火山噴火の後の火山灰の冬の時期に、クリック子音を音素として共有して始まった。おそらくみんなで一緒に歌を歌っているときに、自然発生的に同じ音を出すようになり、音素として確立したのではないだろうか。



図 1 頤の発達により喉頭が降下して、母音を獲得。
[10]

現生人類の身体的特徴は頤(オトガイ)の発達である。舌筋を使ってクリック子音を発声したことで、顎骨が薄く伸びたとみられる。頤の発達した最古の人類化石が見つかったのが、南アフリカのクラシーズ河口洞窟である。1960 年から 30 年以上、対南アフリカ国際学術ボイコットが徹底的に行われたために、この洞窟に興味を示した研究者は少なく、頤の発達を音節の発声と結びつけた研究者は筆者のほかにはいない。

図 1 に示すように、頤が発達すると、口腔底と皮膚の間に空隙が生まれ、肺の気道の出口である喉頭が食道の途中まで降下しても窒息しなくなった。おかげで母音の共鳴が生まれる声道が生まれ(母音を発音できるようになり)、母音アクセントをもつ音節を獲得した。

音節は周波数離散成分である音素と時間軸上の離散成分である拍(mora)をもつ。この 2 つの離散成分を獲得して、語彙数が無限になり、言葉を文法的に接続して複雑なメッセージを送りあうようになった。言語と知能のデジタル化が始まり、言語的人類は生まれた。

3.2 音節の音素性が無限の語彙を生みだす

デンマークの言語学者イエエルムスレウ(Louis Hjelmslev, 189901965)は音節を「アクセントをひとつだけ含む一連の音韻表現」と定義した。[11] アクセントは母音であるので、V, VC, CV, CVC, CCVC, CVCC(C: 子音, V: 母音)はすべて 1 音節に数える。

音素は周波数領域の離散成分である。離散成分とは、パッと耳にすればはっきり違いがわかる成分であり、言語によって音素の組合せ(周波数の区切り方)は違うが、音素を順列して無限の言葉記号(語彙)が得られるところは同じだ。

ゴリラやチンパンジーなどヒト以外の大型霊長類の音声記号は 100 以下であり、ヒトのひとつの言葉(音声)記号が数十万の語彙数をもち、言語は 5000 以上あるので、ひとつの言語で考えると 3 桁(digit)以上、種として比べるとさらに 3 桁くらい桁違いに(100 万倍ほど)語彙数が多い。

3.3 免疫細胞の語彙で対応できた

免疫システムは、鍵と鍵穴のように特異的に結合する抗原と抗体のペアを 1000 万種類以上生み出すことができる。これはひとつの言語に必要なとされる語彙数の百倍ほどだから、音節の獲得による語彙数の爆発的増加に、脳室内免疫細胞ネットワークはまったく問題なく対応できた。

脊椎動物は脳室、脳脊髄液、B リンパ球・T リンパ球をもつ。B リンパ球は神経細胞と同じように外部からの刺激への対応を、二分法と二元論で行い、興奮性と抑制性の信号を送受信できる。B リンパ球は、抗体をネットワーク端末として自由にネットワークできるように進化した神経細胞であると言ってもいいだろう。

感覚器官から脳室に脳脊髄液接触ニューロンが伸びていて、脳室壁に抗原を形成する。大脳皮質上にあるグリア細胞は、海馬で五官の記憶を記録すると同時に、インデックスとして細胞膜上に抗原を提示する。

これら脳室内の免疫細胞ネットワークが脊髄反射を司り、ヒトの言語処理と知能構築の脳内基盤となっていると考えられる。(6.1表2)

3.4 母語の片耳聴覚による脳内の文法処理メカニズム

音節が時間離散成分(拍)をもつおかげで、ひとつひとつの音節が明瞭に聞き取れるため、内容語と文法語を交互に混ぜて発しても、聞き取る側に混乱は生まれない。

ヒトは母語を聞くとき、最初に入った側の耳だけで聞き、反対側の耳からの入力を減衰させている。母語をモノラル聴覚することで、内容語に隣接する文法語の音韻ベクトルを脳幹聴覚神経核の方向定位能力によって解析しているためではないか。

無限の語彙と文法処理が、デジタル第一進化である。

4. 第二進化：文字が文明と概念を生んだ

4.1 文字は広大無辺な土地で発明された

6万6000年前に南アフリカで生まれて世界各地に広がった言語の人類は、6万年以上もの間、文字を持たずに過ごした。文字は、今から5000年前にメソポタミア地方で生まれた。なぜか、メソポタミア地方は、ゴンドワナ大陸がユーラシア大陸と衝突したあと、両大陸の間の海に土砂がたまって生まれた、驚くほど平坦で水利のよい肥沃な土地が延々と続く。南北1000km、東西200~400kmにわたって、丘ひとつない大平原である。

今から9000年前に農耕が始まり、豊かな生産力を背景として王朝が生まれ、人間の感覚能力や記憶能力を超えるほど広大な土地を支配するために、記憶を脳の外で共有するシステムが必要となり、時間がたっても消えない文字は作り出されたのであろう。この仮説を裏づけるように、古代の四大文明はすべて、ゴンドワナ大陸とユーラシア大陸が衝突した際に生まれた大陸間の海(メソポタミア、インダス河平原)か、衝突の衝撃で大陸が割れて生まれた大河の三角州(ナイル河、黄河)で生まれている。

文字は、時空間を支配しようとする王朝の意思と、認識能力を超えた広大な平原の相乗効果で作り出された。伝搬距離が限られ、1秒もせずに消える音声と比べて、時空間に遍在する安定性に文字の真骨頂がある。読み書きを覚えたヒトにとって、文字は、消えない長期的音節(音声)である。

4.2 文字が文明を生んだ

文法は発生も獲得も自然発生的である。これは今も変わらない。これに対して、文字は発生も獲得も作為的である。王朝や政権の都合で文字は政策的に発生したし、一方的に決められる。個人が読み書きを身につけるには学校や塾という教育制度が不可欠である。学校もメソポタミア地方で生まれた。

文字を使いこなすために社会は、学校、図書館、出版業、正書法(音声を文字列にし、文字列を音声にする規則)などを必要とする。これらが整備されると、一定の時空間のなかで言語情報が共有されるようになり、生涯を通じて獲得した知識が次の世代に手渡されるようになった。その結

果、知識や技術が連続的に発展するようになったことが文明ではないだろうか。文字が文明を生んだのだ。実際、文字の使い勝手の良し悪しが、文明の性質に影響を与える。

4.3 低雑音環境で言葉は群に進化した

文字によって文明が生まれ、文明が発展すると、僧院や大学などの世俗から隔離された静かな環境のなかで、不可思議な自然現象や人間存在に疑問を抱いて、昼となく夜となくひたすら学習し思考する人々があらわれた。

彼らは、言葉を厳密に正しく使用することに配慮した結果、言葉の操作によって意味ある解析を行うにいたった。彼らの使う言葉は、数学的な群(概念)に進化したのだ。

それまで言葉は反射的に使われてきた。言葉の刺激は、1対1の二元論理(If A then B)によって、Aという言葉と結びつくある一つの記憶Bを呼びました。もしBが適切でないときは、B'を思いだし、それがだめならB''を思い出すという具合に、言葉はつねに一つの記憶と結びつく。

一方、概念は、1対全の数学的な群の論理が適用される。Aという言葉は、Aに属するすべての元を指し、例外は存在しない。ピアジェは、合成性、可逆性、連合性、同一性、同義性の5つを群性体の条件とした。群であるから概念操作(群の演算)が有意な結果を生み出すのだ。[12]

文明現象と概念操作が、デジタル第二進化である。

5. 第三進化①：物理層の電子化

5.1 原爆とICBMがインターネットの時代を生んだ

ヴァニヴァー・ブッシュ(1890-1974)は電気工学を専門とするアメリカ人の学者で、第一次世界大戦のときに軍と協力して対潜水艦戦の技術を改良した。第二次世界大戦では率先して軍と産業と大学の協力を進めるために科学研究開発局(OSRD: Office of Scientific Research and Development)長官となり、6000人の科学者を率いて戦争のための技術開発を指揮し、原爆開発のマンハッタン計画の責任者でもあった。

発明家であり、アナログコンピュータの開発も行ったブッシュは、戦後の科学者が専門分野の研究を続けつつ、学際的な研究を進めるための記憶拡張装置としてMEMEXが必要と考えた。MEMEXは研究者の机の上に、専門分野に限らない学際的で有用な資料を次々にもってきて、最終的には直接脳内に送りこむ。そして、自分の思考過程を記録したテープを配布すると、他の人たちがそれを共有できる。学者の知能を拡張する夢の記憶拡張システム(MEMory EXtenderでMEMEX)である。[13]

ブッシュのMEMEX構想が雑誌に掲載されたのは1945年7月だった。そのアイデアに予算がついて研究開発が本格化するのには、1957年にソ連が大陸間弾道弾ミサイルの実験に成功し、人工衛星スプートニク(旅人)を打ち上げた後だ。遅れをとったアメリカは国防省高等計画研究局(ARPA)を設立し、人工知能やネットワーク、パソコンにつながる研究開発が始まった。

5.2 軍拡競争のなかで生まれたパソコン

アメリカ人であるダグラス・エンゲルバート(1925-2013)は、1945年8月の日本が敗戦した日にサンフランシスコ港を出港して、1年ほどフィリピンで軍事訓練を受けた。そこでブッシュが MEMEX について書いた記事を読み、人間の知的進化を生み出す装置に心を躍らせた。[14]

エンゲルバートは 1962 年に、人類のために MEMEX を実現しようと拡張研究所(Augmentation Research Center, ARC)で研究開発を指導する提案書を書き、カリフォルニアで研究をつづけた。そこで生まれた技術をもとに、ゼロックスのパロアルト研究所(PARC)で 1973 年にアルトと名づけられたパソコンが生まれる。

現在私達がパーソナルコンピュータで使用しているほとんどすべての要素が、ゼロックスのアルトの中に存在していたのは驚くべきことだ。グラフィカル・ユーザー・インターフェイス付きのディスプレイ、マウス、ネットワーク、電子メール。後に Word と Excel を生み出したチャールズ・シモニーが開発したワープロと表計算ソフト、ラリー・テスラーが考案したコピー&ペースト機能、なんと PDF を考えたチャールズ・ゲシキも PARC にいた。

「ゼロックスは、パーソナルコンピュータを発明したことでも有名にならなかったことで、有名である。」と皮肉られるが、軍事機密が多かったから公表できなかったこともあっただろう。エンゲルバート自身「ラングレーは私たちのスポンサーだった」と語っているし([14]、P221、ラングレーは CIA 本部のある町)、ジョン・マッカーシーの人工知能研究所 SAIL も「世界にも数か所しかない最も優れたコンピュータ・ハッカーが集う場所」と言われていた。[15]

5.3 いよいよ MEMEX は実現に近づいた

1989 年には欧州原子核研究機構(CERN)のティム・バーナーズ=リーがインターネット上の文書提供システムである World Wide Web を開発し、1994 年にはジェフ・ベゾスがアマゾンを創業、翌年オンライン書店がサービスを開始。1998 年にはスタンフォード大学博士課程にいたラリー・ペイジとセルゲイ・ブリンが Google を創業する。キーワードを投入すると、関連する言語情報を教えてくれる検索エンジンの登場で、MEMEX はいよいよ実現に近づいた。

5.4 電子情報は対話する音節

言語情報の電子化が、人類の知能発展にとってどのような意味をもつのかについては、漠然とした議論はあっても、言語発展史的としての議論は少ない。デジタル言語学は、論理成分である音素と拍が、物理層上では音節、文字(時間が経過しても消えない音節)、電子情報(検索に対して対話的に反応する音節)と三段階の進化をしたととらえる。音節しかなかった時代は、直接話ができる相手としか言葉を変えなかった。

文字が手に入ってから、著者は未来や遠方の読者に向けて書き残せるが、後世の読者がその文献に出会うためには縁や運や根気が必要だった。

電子化の時代に入って、読者は検索エンジンや OPAC によって必要とする資料・古本を容易に見つけだせ、宅配や

電子版ダウンロードで容易に入手できるようになった。電子情報をもつ対話性が、言語情報を一方向通信から双方向通信へと進化させたのだ。結果的に現代人は「きわめて大量だが信頼性の保証がない言語情報」にアクセスできるようになった。これをどう活用すればよいのかについて、まだ十分に議論されていない。

	物理進化	物理力	論理進化	適応論理
1	音節	音声(音素と拍)	文法処理 論理概念	二分法と二元論、反射
2	文字と識字	長期的音節	科学概念 概念操作	群論、複雑化
3	電子化	対話的音節	前方誤り訂正 進化	再生、繰返し

(表 1 デジタル言語の三段階進化)

6. 第三進化②：物理層⇒論理層インターフェイス

6.1 言葉はどうすれば脳で記憶されるか

言語情報を、どのように脳内で記憶するかについては、まだ確立された手法がない。そもそも脳内で、どの細胞が言葉の記憶を記録し、どの細胞が五官の記憶を記録するかについても、細胞・分子レベルの仮説はない。

役割	能動・受動	NW 端末	場所	細胞	移動性
刺激受容	賦活	抗原提示凸	脳幹網様体	CSF 接触神経細胞	固定
概念装置	受動/能動	抗体凹/抗原凸	脳脊髄液中	B リンパ球	移動
五官記憶	受動	抗原提示凸	大脳皮質	グリア細胞	固定

(表 2 脳室内免疫細胞によるネットワーク要求解析)

これまでに遭遇したことのない新しい言葉に出会ったとき、我々は聞き取れない。何を聞いたかの記憶も残らない。しかし、その言葉を何度も耳にし、かつ重要であるとわかると、3日ほどでその言葉の音韻波形に対する抗体をもつ B リンパ球が生まれる。(3.3)

2014 年 1 月下旬に「STAP 細胞」がもてはやされたとき、老いも若きもこの新しい言葉の記憶を獲得した。おそらく年齢に関係なく、言葉記憶を生み出すためのスイッチがあるのだろう。ヒトの学習能力は加齢によって減退すると考えられていたが、減退するのは新しいことへの知的好奇心ではないだろうか。

デジタル言語学は、表 2 で示すように脳室内免疫細胞ネットワークが言語処理と知能を司ると考える。言葉の記憶を司る B リンパ球が産生・成熟するためには、① **まず言葉の音韻刺激が与えられ**、② **その言葉の必要性・重要性が認識される必要がある**。

筆者は 2012 年にインドに旅をしたとき、ホテルの名前をうまく言えずにリキシャで間違ったホテルに連れていかれた経験をもつ。その際、正しくホテルに戻るためにはインドのパドマヴァティ女神の名前をリキシャの運転手に伝える必要があると学んだ。そのときは記憶できるわけがな

いと思ったのだが、なんと三日ほどすると、筆者の脳内にパドマヴァティという言葉の記憶が生まれたのだった。

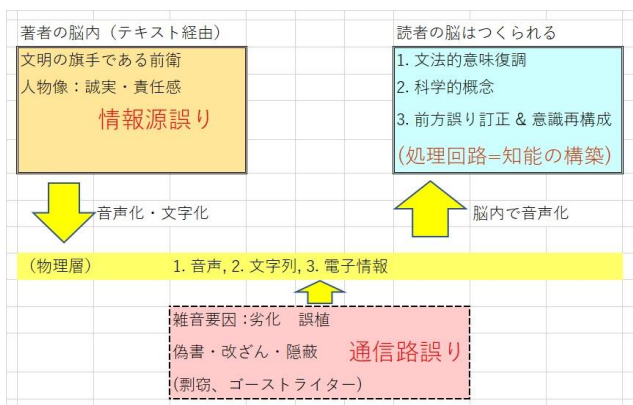
生まれたばかりの B リンパ球は、言葉の記憶だけであり、まだ意味の記憶と結びついていない。どうすれば正しい意味を獲得できるのだろうか。

6.2 言語情報の誤りを 2 種類に切り分ける

「きわめて大量だが信頼性の保証がない言語情報」が我々のものに届くとき、言語情報の信頼性を確かめ、高める手段はあるだろうか。

情報理論は、情報に誤りが発生する場が、通信回線（物理層）か処理回路（論理層）かで、通信路誤りと情報源誤りの 2 種類の誤りに切り分け、それぞれの誤りの性質に適した誤り検出・誤り訂正を行う。

言語情報において、情報源誤りとは著者の責任で発生する誤りであり、通信路誤りは著者の責任に帰すことのできない誤りである。この 2 種の誤りは、A と (1-A) の排中律の関係にあるので、2 つとも誤りを訂正すると誤りを全部正したことになる。



(図 2 言語情報は 2 種類の誤りを内包している)

まず、① (人物像) 誰がその言葉に責任をもっているのか調べ、その人が信頼できることを確認する。(6.3) 次に、② (通信路誤り訂正) その言葉がその人本人の言葉であることを確認する。(6.4) そして、③ (情報源誤り訂正) その人が実験や考察した過程を段階的に追いかけて、その言葉が正しいかどうか確認する。(7.3, 7.4)

こうすると仮に誤りを含んでいたとしても、その誤りがどんな誤りかが特定でき、情報源誤りを訂正することができる。

この誤り訂正方法において、どこの誰が書いたのかわからない情報の信頼性は低く、高める手立てはない。いわゆる流言飛語は信じようがないし、相手にしない。そもそも根拠のないうわさ話を、受け売りする行為は無責任であり、受け売りする人の信頼性は低い。ゴーストライター作品や剽窃も、言葉への責任がぼけるので受け入れない。

6.3 人物像に迫る

本言葉の信頼性を確かめるにあたり、著者の人間性や人物像を把握しておくことは重要である。その人が正直で誠実で、自分の言葉に対して責任をもっているときに、そ

の人になりきって著作を読むことの意味がある。それが確かめられなかったら、その人の著作は読まないほうがよい。

その人が書いた本や論文のほかに以下のようなものが参考になるだろう。個々の文書や資料が貴重な証言をすることもあれば、それらがパズルのピースとなって、より大きな情報が浮き彫りになることもある。

- ① 伝記や自伝、インタビュー、オーラルヒストリー
- ② 手稿 (マニユスクリプト) や関連書類。
- ③ 家族や恋人、友人、弟子が書いたその人の思い出
- ④ その人の蔵書 (大学図書館の〇〇文庫)
- ⑤ その人が住んでいた家の訪問や町の散策、墓参り

6.4 通信路誤り訂正

著者の言葉を言語情報として取り込むとき、それが著者本人の言葉、著者が自分で考えた言葉であることを確かめる必要がある。

著者存命中に著者校正された印刷物 (PDF 含む) は、真筆とみなしてよい。時間的劣化、誤植・校正漏れなどは、冗長性で乗り越えることができる。

著者が亡くなった後に行われた改ざん・偽書は、見分けにくい。その著者のものとされるすべての著作物に最低一回は目を通すと、どれが異質なものであるか見えてくる。

著者は死後に改ざんや偽書があらわれることを予見したら、冗長的な符号を付す誤り訂正符号化によって、自分のテキストを保護することができる。[16]

ゴーストライター作品や剽窃は、誰も手直しや追加説明がなされない。誰も書かれた言葉に責任を取らないから、ゴーストライター作品や剽窃の可能性のあるものは読んではいけない。

7. 第三進化③: 論理層で前方誤り訂正

7.1 科学の誤りが訂正されず放置されている

20 世紀は、量子力学と分子生物学という不可視の領域を解明する科学が隆盛した。1965 年にノーベル生理学・医学賞を受賞したフランス人分子生物学者フランソワ・ジャコブは、最晩年に書いた本のなかで告白する。

「20 世紀という科学的探究の世紀最大の発見は、わたしたちが自然についてじつは何も知らないということであろう。学べば学ぶほど、無知の大きさがわかってきた。このことはそれ自体、ひとつの大いなる報せである。この報知は、18 世紀や 19 世紀の先達を驚かせるに足りるだろう。いまはじめて、わたしたちは自分の無知を目の前に見ている。長い間、わたしたちは事物がどのように機能するかを知っていると称していた。しかしそれは単に、穴を塞ぐためにデタラメな話をしていたようなものである。」[17]

科学は伝統的生物学の誤りを露呈した。しかしその誤った古典的科学的デタラメはまだ正されないまま、教科書に掲載されている。

たとえば脳科学では、言葉の記憶や五官の記憶を記録する分子構造を解明できないまま、古い教科書に掲載されていたことが今もそのまま通用している。大脳皮質のプロカ野とウェルニッケ野で言語が処理されているという話や、脳脊髄液は脳を保護するためにあるという話である。

あるいは、熱力学と情報理論では同じ「エントロピー」概念が逆の意味となるとされる。どうして同じ概念が逆の意味をもちうるのだろうか。エントロピーという概念は、いつ、どこで、誰が名づけた言葉であるのかと考えれば、一つの意味しかもちえず、学際的に通用するはずである。

熱力学の分野では、日本の八木江里博士が、ポーランドの物理学者クラウジウス(Rudolf J.E. Clausius, 1822-1888)が最初の論文集に収録した 17 論文から全ての方方程式(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。[18]「構成要素の配置の変換値(エントロピー)」である。

この概念は通信理論でも有効である。シャノンの一般通信モデルには、通信回線上に雑音源(noise source)が描かれている。これは太陽からの放射熱や降雨の影響で、情報を伝える電波の伝播が阻害されることをうまく説明する。情報理論は、デジタルな通信理論であると考えれば、そのエントロピー概念も、熱力学と同じではないか。

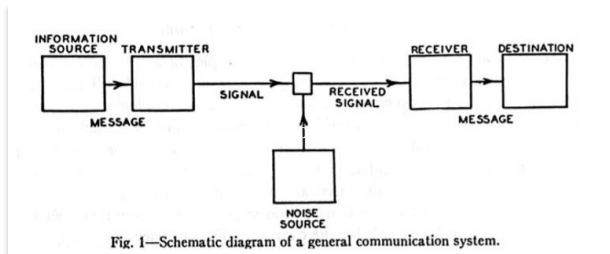


Fig. 1—Schematic diagram of a general communication system.

(図 3 シャノンの一般通信モデル) [19]

情報理論が、エントロピーを熱力学ではないというのは、学際的研究やシャノン概念の起源研究が進んでいないために、誤りと混乱が放置されているだけではないだろうか。

7.2 言葉の意味の複雑次数

言葉の意味は重要だが、言語学で意味論を扱う学者は少ない。「意味はあるコトバと結びついた個人の記憶(の総体)」という鈴木孝夫の意味論は明解だ。[20]

言葉の意味は個人の記憶だから、たとえば「クスクス」という料理を食べたことがない人は、その言葉の意味をもたない。その人が実際に食べたとき、味わった五官の記憶が意味を構成する。(複雑次数 1)

鈴木 の 定 義 で は、五 官 の 記 憶 を 意 味 と す る 概 念 を 二 元 的 に 操 作 し て 生 ま れ た 思 考 の 記 憶 も 意 味 と なる。筆 者 は そ れ を 一 次 論 理 概 念 と 呼 ぶ。(複 雑 次 数 2) そ れ は た と え ば :

類 の 概 念 : 「 家 具 」 = 「 机 OR 食 器 棚 」, 「 柑 橘 類 」 = 「 レ モ ン OR 甘 夏 OR 金 柑 」,

関 係 性 の 概 念 : 「 首 都 」 = 「 中 国 AND 北 京, 日 本 AND 東 京, フ ラ ン ス AND パ リ 」, 「 幼 虫 」 = 「 ア ゲ ハ 蝶 AND 青 虫, ト ン ボ AND ヤ ゴ, 蚊 AND ボ ウ フ ラ 」,

こ れ ら は 具 体 的 な 概 念 を OR や AND と い う 単 純 な 式 で 操 作 し て 生 ま れ る 記 憶 の 集 合 に 名 前 を 与 え た も の だ。

一 方 「 STAP 細 胞 」 の よ う な 科 学 的 概 念 (高 次 論 理 概 念) も 思 考 記 憶 の 積 み 重 ね て 生 ま れ る。た だ 積 み 重 ね る 言 語 が、「 B リ ン パ 球, 活 性 化, 子 宮, 脾 臓, 抗 体 」 の よ う に 五 官 で 感 じ ら れ ず, 事 前 に 学 習 を 必 要 と す る 科 学 的 概 念 で あ る こ と と, 積 み 重 ね 方 (論 理 の 組 み 立 て 方) が AND や OR の よ う な 単 純 な 二 元 論 に と ど ま ら な い と こ ろ に 大 き な 違 い

が あり, 複 雑 度 が 高 く なる。(複 雑 次 数 3)

概 念 の 具 象 ・ 抽 象 性, 複 雑 さ の 次 元 の 違 い は 重 要 で あ り, そ れ を 感 じ 取 る た め に は, 複 雑 性 に い く つ か の 次 数 が あ る こ と を 認 識 し て お く 必 要 が あ る。

複 雑 次 数 0 : 脊 髄 反 射 の よ う に 直 接 行 動 を 必 要 と す る 言 語 (If A then B: 「 津 波 」 ⇒ 「 高 い と こ ろ に 逃 げ る 」) や 固 有 名 詞 (同 姓 同 名 の 別 人 物 は お 呼 び で な い) に は 一 般 化 ・ 総 合 化 の 要 素 が な い。

複 雑 次 数 1: 五 官 の 経 験 を 積 み 重 ね る こ と に よ っ て 意 味 を 深 め 広 げ て 総 合 す る 一 般 概 念 (「 机 」, 「 温 泉 」, 「 ト マ ト 」 な ど)

複 雑 次 数 2: 二 つ の 一 般 概 念 を 単 純 な AND や OR の 操 作 に 投 入 し て 生 ま れ た 思 考 結 果 の 記 憶 の 集 合 を 意 味 と す る 類 や 関 係 性 の 概 念 (一 次 論 理 概 念 : 「 柑 橘 類 」, 「 首 都 」, 「 幼 虫 」 な ど)

複 雑 次 数 3: 複 雑 な 現 象 に 名 前 を 与 え た が 学 際 的 に 通 用 す る 概 念 に な っ て い な い。分 野 科 学 内 だ け で 通 用 す る 概 念 (「 本 能 」 「 言 語 脳 」 な ど)

複 雑 次 数 4: 思 考 記 憶 の ネット ワーク に よ っ て 意 味 を 構 成 し, 学 際 的 に 共 有 で き る 明 解 な 定 義 を 有 し, 数 学 的 な 群 の 要 件 を み た す こ と に よ っ て 概 念 操 作 の 結 果 が 有 意 と なる 科 学 的 概 念 (「 (真 核 生 物 の 核 内 で お き る) 転 写 後 修 飾 」, 「 エ ン ト ロ ピ ー 」, 「 遺 伝 」 な ど)

言 語 や 概 念 に 複 雑 次 数 が あ る こ と を 共 通 理 解 と し て, 子 供 た ち に も 教 え る べ き で は な い だ ろ う か。英 単 語 を 使 う と き に 加 算 か 不 加 算 か を 意 識 す る よ う に, 言 語 を 学 ぶ と き 使 う と き に 常 に そ の 言 語 の 複 雑 次 数 を 念 頭 に お け ば, 学 習 は 容 易 に な り, 誤 解 は 減 り, 人 類 の 知 的 水 準 は 向 上 す る だ ろ う。

英 和 辞 書 で は 「 不 」 や 「 U 」 の 記 号 で 不 加 算 名 詞 (複 数 形 を つ く ら な い) を 表 示 す る よ う に, 国 語 辞 書 や 英 和 辞 書 で そ れ ぞ れ の 名 詞 を 「 名 詞 (複 0) 」, 「 noun (comp4) 」 と い っ た 具 合 に 0 か ら 4 の 数 字 で 概 念 の 複 雑 次 数 を 表 示 す れ ば, 複 雑 次 数 の 考 え 方 は 早 く 広 ま る だ ろ う。

こ れ ま で 複 4 だ と 思 っ て い た 概 念 が, 実 は 複 3 で あ る こ と に 気 づ く だ け で, 科 学 は 混 乱 を 回 避 し て 発 展 す る だ ろ う。た と え ば 「 STAP 細 胞 」 は ま だ 複 3 の 段 階 に あ り, そ れ が 複 4 に な る か を 論 文 で 世 に 問 う た の で は な か っ た か。ま た 「 エ ン ト ロ ピ ー 」 は 複 3 と し て 使 う べ き で な く, 複 4 の 学 際 的 概 念 と し て 使 う べ き で は な い だ ろ う か。

複雑次数	言葉の種別	例
複 0	災害時の行動 固有名詞	津波, 火事 特定の人物の姓名
複 1	具象概念	電車, 温泉, トマト
複 2	一次論理概念	柑橘類, 首都, 幼虫
複 3	複雑概念	仏向上, 本能, 言語脳
複 4	科学的概念 (学際的・群)	遺伝, エントロピー

(表 3 言葉の複雑次数区分の例)

7.3 科学的概念は個人の知的発見として取り扱う

新しい科学的概念は、それまでの言葉で説明できない現象や物質を発見した科学者が、それにふさわしい名前をつけることで生まれる。

だとすれば、誰が、いつ、何を、どのようにして発見し、それをどう名づけたかを、著作や論文を通じて確かめることが科学的概念を取り扱ううえで必要な作業ということになる。それは知的所有権である特許の出願と審査のような手続きであり、複雑回数 4 の科学的概念として適切であるかどうかを審査することが求められる。科学的概念は著作物に相当し、その発見者は著作権者に準じた保護がなされるべきであろう。

たとえば 2014 年 1 月末から 2 月にかけて日本のテレビ報道は、STAP 細胞をもてはやしたかと思えば、すぐに有無をめぐって大騒動した。その際、Nature 論文の第一著者を質問攻めにしたのは過ちであり、論文を精読するべきではなかったか。[21]

2 月頭に筆者は Nature の記事を読み、マウスの B リンパ球に刺激を与えて、その B リンパ球を別のマウスの子宮内で培養すると、3 日間活性化した現象を刺激惹起性多能性獲得細胞(Stimulus-triggered acquisition of pluripotency)と名づけたことを理解した。しかしこの活性化は、利根川進博士がノーベル講演のなかで、免疫二次応答において新しい抗原に抗体を最適化するための B リンパ球の活性化、体細胞超変異(Somatic Hyper-mutation)と似ている。もしかするとこれは STAP 細胞と呼ぶよりも、B リンパ球の STAP 化(活性化)を誘発したと考えたほうが妥当ではないかと思ひ、議論が広がることを期待していた。その後の展開はまことに残念であった。

概念は特定の学問分野に限られるものではなく、学際的・普遍的な現象である。しかしそれは低雑音環境で切実精思して得られたものであり、きちんと論文を読んでいないワイドショーのレポーターがずかずか足を踏み入れることが許される領域ではない。その概念を理解したければ、論文を読んで著者の発見ドラマを仮想現実的に読み解くほかない。

7.4 生命体のデジタルな自己増殖に学ぶ

多細胞真核生物は、受精卵核内の DNA の発現制御と細胞間ネットワーク制御によって、細胞分裂をくり返すことで、親と同じだけ複雑か、もしくは親よりも進化した子を生みだす。

遺伝情報は親から子に送られるだけで、子から親への遺伝子レベルのフィードバックはない。一個の受精卵から一個の生命体への分裂が始まる仕組みは、誤りを生みにくく、環境条件が変化するときなど突然変異を誘発して、新しい条件に最適化して進化しやすい。生命体の生殖は、デジタル通信であり、前方誤り訂正が行われている例といえよう。

ヒトの知能において複雑さは、言葉記号を獲得してから、その言葉の複雑回数に応じた意味を正しく構築するところにある。(7.2) 誰かの概念を獲得するにあたって、受精卵同様に、きわめて基本的なところ(名前の記憶)から、段階的に丁寧に思考操作を重ねることで、意味を複雑化していくとよいのではないか。

著者は、事実を歪めることなくありのままに記述する。できるだけ正確に書き残す。書くという行為は、人類共有知への自己の貢献であると理解する。こういったことが著者に求められる。

読者には、本や論文に書かれた言語情報から、直接体験することなく、現実におきたことを正確に読み解く技量が求められる。教養とは、言語情報から現実を復元して知識や知恵を得ることではないか。最近、大学で教養課程が短縮化されたことは、これから学際的統合をする研究が増える科学にとってはマイナスである。

読者は、著者の人物像を理解し、著者本人の言葉であることを確認したら、著者になりきって、言語情報の背景にある実験結果と観察と思考とを、自分の意識内の現象として受け入れるのだ。一度や二度の読書では、なかなかそこまではできないから、時間をおいてくり返し、読む速度や集中力を変えながら、著作と一体化する知識再生型(regenerative)読書が必要である。

著者よりも後の時代を生きる読者は、著者がもっていなかった科学的知見や電子的ツール(検索、ダウンロード、複写、電子メール、zoom 会議など)をもっているのだから、著者には解明できなかったことを解明できる可能性がおおいにある。こうして知識や概念の前方誤り訂正と進化が実現するのである。

巷間の書籍をみると、多くの著者が別の学問分野の著作で読んだ知識や概念を丸のみして、吟味や確かめや再生的受容をしないまま、安易に受け売りしていることに気づく。ここに科学の混乱の元凶がある。丸のみや丸暗記というアナログな学習法では、知識や概念に含まれている雑音や歪を除去できずに、そのままそれを増幅して次の世代に伝えることになってしまう。

知識再生型読書と前方誤り訂正の手法を確立すれば、人類共有知性は、大きく進歩することになるだろう。

8. おわりに：サイバー世界で会いましょう

8.1 知能の自己増殖過程に突入した

言語的人類は、洞窟のなかで喉頭降下によって音節を獲得し、無限の語彙と文法を手にした。続いて、ヒトの認知能力を超える広大な平原で文字を共有して文明を生みだし、静かな僧院のなかで科学的概念を得た。これが人類知能の第一進化と第二進化であった。

今、電子情報を得たおかげで、キーワード検索とワールドワイドウェブへのアクセスができ、有史以来の人類知能の前衛たちと出会えるようになり、また世界のどこに住んでいようとサイバー空間を通じて同時代の研究者との研究会や学会に参加できるようになった。いよいよ知能進化はデジタルな自己増殖(オートマトン)過程に入ったのだ。コロナウィルスのおかげでオンライン開催の学会・研究会が増えたことは、知能進化にとっての追い風であった。

論理層構築において重要であるのは、超低雑音環境にして、概念の誕生経緯を読書で追体験することだ。また自分が再生的に受容していない概念や理解不十分な概念は使わないようにすることだ。

8.2 前衛を知る喜びと前衛に立つ喜び

孟子の言葉に「天下の善士を友とするをもって未だ足らずと為せば、また古(いにしえ)の人と尚論(しょうろん)す。」(世界級の人を友としても足りないなら、書物を通じて古人と語り合う)とある。

デジタル言語学は、霊長類学者島泰三の「はだかの起原」に触発されて始まった。[22] それから、すでにこの世にいない発達心理学者 ピアジェ、脳生理学者 パブロフ、数学者 ノイマン、免疫学者 イェルネ、心理学者 ヴィゴツキー、動物生態学者 ティンバーゲン、進化生物学者 メイナードスミス、言語学者 千葉勉、脳生理学者 時実利彦ら、卓越した人類知能の前衛とよぶべき研究者の仕事を読むことで発展した。

図書館 OPAC や検索エンジンや読んだ論文の参考文献リストを通じて、人類知能の前衛と呼ぶべき学者たちの著作に出会えたことは、筆者のこれまでの不勉強と無知を恥じることにつながったが、一方で彼らの著作から学び発見したことが大なる喜びを与えてくれた。

これまでいくつか学際的な文献を読んだ結果、情報には、(i) 卓越した研究者が広い視野で全体像とその時点の学問の到達点を描いたレビュー論文である高次情報、(ii) 誠実な研究者が書き残した一部誤りもあるが多くは正しい、情報源の誤り訂正が必要な情報、(iii) 間違っているだけでなく読者の思考枠組みを歪めかねない情報、と少なくとも3種類あると思った。

残念なことにマスコミや学会がもてはやするのは、(iii)の学者が多く、多数派の意見を鵜呑みにして、間違った情報いいかげんな似非科学ばかりがさらに広まっていくのは、残念なことである。信頼性のある情報にたどり着くには、(i)や(ii)の学者を知り、彼らの著作から業績を知り、時間をかけて自分の頭のなかに取り入れる努力が必要だ。彼らが興味を示したこと、やろうとしてやり残した仕事は宝の山である。著者が十分な説明をしないまま、未熟な発想を書き残している場合、読者はそれがなんだかわからない状態のまま、頭の隅にとどめておくのがよい。それこそが宝であり、いつかヒラメキを生みだして新たな発見に結びつくかもしれない。意外なほどに、この宝の山に気がついている人は少ない。

電子情報を上手く利用するコツは、システムを物理層と論理層に分けて、物理層は良い本や真摯な研究をした人と出会うところまでと割り切って、その先の知識獲得や誤り訂正は超低雑音環境に引きこもって、知識再生的読書を実践することで、脳内論理層に前衛たちの仕事を取り込み、自らが前衛に立つことだ。

ヒトは学ぶサルである。今を生きる者だけがこれまでの人類の知的営為を学び、誤りを訂正し、進化させることができるので、その重大な責任を理解する必要がある。研究成果はオンラインの学会・研究会で発表するほか、論文にしてクリエイティブコモンや CiNii、情報学図書館で保管してもらえば、きっと将来誰かがみつけてくれるだろう。

謝辞

本研究は2009年10月の電子情報通信学会思考と言語研究会ではじめて研究会発表をして以来、情報処理学会と電子情報通信学会、人工知能学会に属する研究会で発表してきたことの成果です。拙い発表で、毛色の違ったテーマであったにもかかわらず受け入れてくださった研究会の幹事ならびに参加者の皆様に様々なご迷惑をおかけしたことをお詫び申し上げますとともに、心から感謝申し上げます。

参考文献

- [1] K. Tokumaru, Inside-Cave Evolution Hypothesis of Modern Human. (OP4-15), 10th JSES (2008) 日本進化学会第10回東京大会
- [2] Pepper W.J., Braude S.H., Lacey E.A., Sherman P.W. "Vocalization of the Naked Mole-Rat", in Alexander RD et al(1991) The Biology of the Naked Mole-Rat, Princeton Univ.
- [3] 得丸 記号と言語と条件反射～鈴木孝夫の『鳥類の音声活動—記号論的考察』に刺激されて 信学技報 TL2012-1
- [4] Noll, H. The Digital Origin of Human Language - a Synthesis, BioEssays 25:489-500, 2003
- [5] 得丸 デジタルな言語記憶に関する仮説 情処学会研究報告 2011-NL-200 No.1
- [6] 得丸 モバイル・ニューロン(1)～パブロフはなぜ犬の大脳皮質をくり返し切除したのか 信学技報 NC2020 (印刷中)
- [7] 得丸 自然言語の分子構造～意識の脳室内免疫応答仮説, 信学技報 NLC2012-1
- [8] 得丸 母語のモノラル聴覚と文法処理-例外としてのビダハン, 情処研究報告 2014NL-219No.23
- [9] メイナードスミス, サトマーリ 生命進化 8つの謎, 朝日新聞社 2011
- [10] 千葉・梶山, 母音—その性質と構造, 岩波書店 2003
- [11] 得丸 物理的音節と論理的音節—言語ネットワークの量子メカニズム, 情処学会研究会報告 2012-SLP-91-No.16
- [12] ピアジェ, 知能の心理学, みすず書房, 1965
- [13] ブッシュ 考えてみるに, 1945, 山形浩生訳
- [14] バーディニー, ブートストラッパー人間の知的進化を目指して ダグラス・エンゲルバート, あるいは知られざるコンピュータ研究の先駆者たち, コンピュータエージ社, 2002
- [15] マルコフ パソコン創世「第3の神話」カウンターカルチャーが育んだ夢, NTT 出版, 2007
- [16] 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 COMP2018-9
- [17] ジャコブ ハエ, マウス, ヒト 一生物学者による未来への証言, みすず書房 2000
- [18] クラウジウス, 八木江里監訳, エントロピーの起源としての力学的熱理論 クラウジウス熱理論 論文集, 東海大学出版部, 2013
- [19] シヤノン, ウィーバー 通信の数学的理論, 植松友彦訳, ちくま学芸文庫 2009
- [20] 鈴木孝夫 ことばと文化, 岩波書店, 1975
- [21] 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知する, 信学技報 SeMI2020- in printing
- [22] 島泰三 はだかの起原, 講談社学術文庫, 2018