

ヒトの音声は phonit でデジタル変調されている (復調はアナログ方式である) - デジタル言語学(その 1) -

得丸公明(衛星システム・エンジニア)

〒158-0081 東京都世田谷区深沢 2-6-15

深刻化する地球環境問題について考えていた筆者は、人類とは何か、文明とは何かという根源的疑問に出会った。2007年に人類発生の起源を訪ねて最古の現生人類遺跡である南アフリカのクラシーズ河口洞窟を訪問したことをきっかけに言語の起源に関心をもち、情報理論の一般通信モデルを援用して考察した結果、言語はデジタルシステムを構成していることに気づいた。人類と他の動物の唯一・最大の違いは、人類は音声通信をデジタルシステムへと高めたこと、離散・有限信号である音節を使って、順列組合せにより符号語の数を無限に作れるようになったこと、一、二音節のわずかな音韻の変化・追加によって、概念を接続・修飾する文法を生み出したことにある。2009年10月以来、試行錯誤しつつ考えてきたことを情報処理学会や電子情報通信学会・人工知能学会の研究会で発表してきた。(1)(2)(3)(4)(5)(6)(7)

今回は音韻学における古典的な問題である、「発話された音素の伝送レートが、聴覚の処理限界を超えている」事実に着目し、ヒトの言語では変調はデジタルであるが、復調はアナログに行なわれていることについて論ずる。

Human Voice is Digitally Modulated with "phonit" and Analogously Demodulated. - Digital Linguistics (Part-1)

Kimiaki Tokumaru (Satellite System Engineer)

2-6-15, Fukasawa, Setagaya-ku, Tokyo 158-0081 Japan

When the author was contemplating on the aggravation of Global Environmental Problems, he became interested in the origins of humans and their civilizations. In 2007, he made a visit to the Klasies River Mouth Caves in South Africa, the oldest modern human site in the world, and realized that human and language were born inside caves. Exploiting the Information Theories, he came to the conclusion that human language system is digital in applying discrete combinatorics to create infinite number of phonetic symbols and in serialized, complex and long message generation combining conceptual and grammatical words. (1)(2)(3)(4)(5)(6)(7)

In this paper, a phonological phenomenon that "the phoneme transmission rate in

human speech is beyond human's audition ability of phoneme recognition", will be investigated. The author hypothesizes that human speech is digitally modulated but demodulated analogously.

1. はじめに

1.1 ヒトの音声通信はデジタル通信である

1.1.1 通信回線の音声的な多様性

動物が仲間同士の間の音声通信で用いる信号の数は数十種類であるのに、ヒトは数万、数十万もの単語を使いこなす。大気振動として瞬間的に構築される音声回線上の信号特性に注目すると、動物は感情を反映した周波数変移パターンである同じ信号を何度でも繰り返すのに、ヒトは一定数の音節を組み替えて長く複雑なメッセージをつくり一度送るのみである。ヒトは、離散信号(音節)セットを言語共同体成員間で共有し、生後に離散的発声と離散的聴覚を身につけるから、送信メッセージは劣化や歪みを伴わずに忠実に相手に届く。これはデジタル通信の特徴である。

1.1.2 概念の数の多さ

日本語は、平安時代以来、ひらがな・カタカナという世界でも珍しい音節文字を2組も用いてきた。音節はかならず母音で終るし、音韻論上で一定の時間長を持った基本単位「拍(モーラ)」をベースに発音する。(ちなみに長音「ー」、促音「っ」、撥音「ん」は1モーラとして規定され、拗音「ゃ」「ゅ」「ょ」は1モーラとならない。)

たとえば、2モーラの「あん」(案、餡、庵)自体ひとつの単語であるが、その末尾に1モーラ加えただけでも「安易」、「あんか(安価、行火)」、「あんき(暗記、暗鬼)」、「行脚」、「暗渠」、「暗愚」、「餡子」、「安居」、「暗示」、「杏」、「あんた」、「あんだ(安打、編んだ)」、「あんち(アンチ、安置)」、「あんど(アンド、安堵)」、「あんな」、「暗に」、「鞍馬」、「安否」、「暗譜」、「按摩」、「暗喩」、「あんよ」などたくさん単語が思い浮かぶ。「ん」の撥音を省いて2モーラにすると「愛」、「赤」、「秋」、「顎」、「味」、「仇」、「穴」、「兄」、「尼」、「鮎」、「ん」を長音や促音に変えれば、「アーク」、「アーチ」、「アート」、「あっか(悪化・悪貨)」、「悪鬼」、「あっち」などの単語になる。

音節文字を組み合わせれば、いくらでもたくさんの単語を作り出すことができる。離散信号の組合せによって無限の数の単語を作り出すのも、デジタル原理である。

1.1.3 概念を結合・修飾する法則・規範としての文法

送信者が発信したメッセージが、一度だけの送信で、一音たりとも誤たずに受信者に伝わるようになると、わずかな音韻の変化や追加によって概念を接続し修飾し演算する法則語が生み出された。この概念を接続・修飾・演算するための共有された文章規範である法則語を文法と呼ぶことにする。概念は個人の知覚にもとづく記憶であり

自立語であるが、文法は具体的な事物や存在の記憶とは結びつかず、単独で用いられることもない付属語である。名詞、形容詞、副詞、動詞は概念であるが、接続詞、関係詞、疑問詞、感嘆詞、代名詞、助詞、助動詞、動詞と形容詞の活用語尾、数詞、冠詞、敬語、性、限定詞は文法である。文法のおかげで、時制、仮定・条件、意志などきわめて繊細な意味内容をメッセージに付加できる。

こうして概念と文法を組み合わせ、長くて複雑なメッセージを順序だてて送信することが可能になった。これが人類文明発展の原動力となった。

1.2 デジタル通信に固有な通信誤り

母語であっても、同じ単語が違った意味に受け取られる意味誤り(情報源符号化誤り, semantic error)は、同音異義語(例:水田の土壌か泥鰌か)による勘違いや一休さんのとんち話の「このはし渡るべからず」「橋」と「端」の意図的な解釈誤り)のように枚挙にいとまがない。

一方で母語を共有する者同士で行なう通信では、音素や文法が共有されているため、音素誤り(phonemic error)や統語誤り(syntax error)には出くわさない。異文化交流で発生した音素誤りや語順誤りの例を、以下で紹介する。わずかな音素や語順の乱れが著しい誤解を生んだ実例は、言語がきわめて精巧で繊細な通信であることを示す。

1.2.1 伝送路符号化誤り

日本語にない音素は、聞き取ることも発音することも難しい。フランスに駐在したばかりの日本人主婦が、近所の食料品店で買い物をしたときの話。支払いのレジで、

店員: C'est tout? (お買い物はこれで全部ですか?)

日本人客: セトゥー(店員の台詞をオウム返ししたつもり)

店員: Sept oeufs? (卵を7つですか?)

オウム返ししたつもりだったのに、c を s として発音したために店員は間違っ理解した。この日本人は、それ以上口を開いて余計なものを買わないように、店員がわざわざ卵を7つ包んでくれたものをありがたく買って帰ったという。

音素セットを異にする言語間の会話においてこの種の誤解はしばしば発生する。それは音素誤り(phonemic error, 音韻誤り phonetic error)、離散性不適合、変調誤り、復調時の誤り訂正機能不全として考えることも可能である。母語にない音素はそもそも認識できないので、話者が気をつけるだけでは不十分で、聞き手も外国人の発音は補正しつつ聞くよう心がけるべきであろう。

1.2.2 通信路符号化誤り

文法誤り(syntax error, 語順誤り)という言葉は、コンピュータが期待通りの作業をできなかったときに人間に送り返すメッセージとしてなじみがあるが、それは人間間の

通信においても起きる。以下はパリでタクシー運転手の体験談として聞いた話。

アメリカ人観光客が乗り込んできて"Eiffel Tour"(トゥール・エッフェル)といった。運転手は"Et fais le tour."(一回回って)と聞き取り、客を乗せた場所で車をグルリと旋回させた。

語順の間違い(正しくは Tour Eiffel)が意味の勘違いを生んだ。いっそ「エッフェル・タワー」と英語読みすれば間違いは生まれなかったかもしれない。シンタックス・エラーが起こりえるのは、話し言葉がデジタル通信だからではないか。

1.3 ヒトと動物はデジタル通信かアナログ通信かの違い

1.3.1 ヒトと動物の音声通信の違い

他の動物の音声通信は同じ信号を何度も繰り返すのみで(ヒトでも幼児は泣き声を繰り返す)、音韻を微妙に変えて長く複雑な文章を組み立てている様子はみられない。ウィナーは「人間機械論」の中で以下のように指摘する。

「人間の通信を他の大部分の動物の通信と区別する特徴は、(a) 使われる符号体系の精巧さと複雑さ、(b) この符号体系の高度の任意性である。多くの動物は自分たちの情緒を相互に信号で伝えることができ、そのさいのそれらの情緒は敵の存在とか、同じ種の異性の個体の存在とか、こういう種類のきわめてさまざまな詳しい情報を表示する。これらの通報の大部分はその時かぎりのもので貯蔵されない。そのかなりの部分は人間の言語に翻訳すれば間投詞や感嘆詞になるものだろう。ただし、一部は名詞と形容詞のような形の語として大まかに表現できるかもしれないが、それらの語を当の動物は人間の言語の場合のような文法的な形の区別なしに使うのである。一般に、動物の言語は第一には情緒を、第二に事物を伝え、事物の間のもっと複雑な関係は全く伝えないように思われる。」(8)

今日携帯電話、テレビ放送、無線 LAN などの分野で利用されているデジタル通信と同様に、ヒトの音声通信も(i) 回線上で離散・有限信号の一次元配列として音節が送受信されているほか、(ii) 生後に発達する脳内言語処理神経回路上で、(a) 単語と意味(記憶・法則)を結びつける情報源符号化、(b) 脳内回路上の無音の言葉である「内言(inner speech)」を大気の振動(音波)に変換する伝送路符号化が行なわれている。(c) 概念と文法を組合せて聞き取りやすい文章にすること、自然の音表象を語源として単語の音韻構造に秩序をもたせエントロピーを低下して誤りをおきにくくすることは、通信路符号化である。

動物の音声通信は、上の(i)の特徴も(ii)の特徴も示さない。言語は、デジタル通信に必須である伝送路・通信路・情報源の3種類の符号化を行っている。これに対して、動物の音声通信を便宜上「アナログ通信」と呼び、両者を対比してとらえるのは、有効であろう。

1.3.2 概念と文法を組み立てて可能となったシリアル通信

昔話には、「昔々あるところにおじいさんとおばあさんが住んでおりました」から始まって、「めでたし、めでたし」で終わるものが多いが、これらは物語の始まりと終わりを特定するための符号で、二つの符号の間で物語が展開することを示す。これらの符号語は電子情報のシリアル通信におけるプレアンブル(開始語)とポストアンブル(終始語)に対応し、あるいは遺伝子の通信文メッセンジャーRNA において翻訳の開始・完了を指示する開始コドン・終止コドンとも対応する。

ヒトと動物の違いの根本は、音声通信がデジタル方式かアナログ方式かの違いであり、ヒトはそれによって無限の造語能力を手にし、長く複雑なメッセージを送りあうシリアル通信ができるようになった。それ以外の部分はヒトと動物は肉体的にも精神的にも同じと考えられる。動物も心や感情をもつ。言葉をもたないから表現できないだけなのだ。

1.3.3 多元音声デジタル信号フォニット

一般的にデジタルという言葉からは、デジタルコンピュータで採用されている二元デジタル信号ビット(binary digit)を想像する。ナノ秒単位の高速度処理は可能だが、微妙な違いを判断することができない機械装置の場合には0か1かの単純な二元デジタル信号が適している。ヒトや動物の中枢神経はナノ秒に比べると100万倍規模で処理速度が遅いミリ秒単位の処理であるが、一方で繊細な感覚器官と中枢神経ネットワークの並列処理のおかげで音素や音階のわずかな違いを識別できる。ヒト言語はビット信号を使用していないことを強調するために、言語共同体ごとに一定数存在する離散的音声信号を、フォニット(phonit = phonetic digit)と呼ぶことにする。

2. 変調レートと復調能力の差

本稿では、伝送路符号化・復号化過程、すなわちデジタル変復調過程について論ずるつもりである。音素の発声は、ヒトに固有な発声器官の発達と発声器官運動制御によって、有声化した肺気流をフォニットで変調することが明白であるのに対して、聴覚器官から大脳新皮質一次聴覚野に送り込まれる音声がいっただいどのように処理されているのかはまだ解明されていないので、復調の問題について考える。

離散的発声能力と離散的聴覚能力の間に伝送レート上のギャップがあることは Liberman らが指摘している。(9, pp432-3, 拙訳)

「音響アルファベットにとまなう困難のうちで、最大のものはレートである。会話においては、最大限一分間に300単語の理解が可能である。それぞれの英単語が平均4から5の音素であるとする、一秒間に30の音素が生み出されていることになる。し

かしながら、聴覚心理物理学の研究によれば30音素/秒というのは耳の時間分解能力を超えている。そのレートでは離散的音響イベントは分析不能なひとつのうなりとして融合してしまう。聴覚者はそのうなりの音程から、話者がどれくらい速くしゃべっているのか判断できるかもしれないが、何がしゃべられたかはほとんど認識できない。通常の会話でめずらしくない15音素/秒であったとしても、耳は音素を離散的な音響イベント列として取り扱う能力をもたない。

音響アルファベットが達成できない要求事項は少なくとももう一つある。識別可能な音響の数が十分であるかということである。音素の数、つまり音響的な形状の種類は何十もある。英語の場合、およそ40である。もちろん、メロディーを聴くときのように、時間とともに刺激をパターン化すれば40の識別可能な音響を見出すことができるだろう。しかし我々が話す速度では、音素セグメントは平均して50ミリ秒以上持続することはない。聴覚認識においてそのような短い持続時間の刺激を正確に識別できるのかということは明らかにされていないが、入手可能なデータをみるかぎりその数は40よりもかなり少ないようである。

したがって、信号の識別可能性を高めるために、なにかほかの符号化・復号化メカニズムの特徴が考慮されているのかどうかということは興味深いことである。」

音声の音素レートは、通常の会話速度でも、耳の離散的処理能力を超えている。そこで Liberman は、聞き取った音声信号を、発声器官運動制御の神経刺激を参照にして音素識別を行なっているのではないかという「運動理論」仮説を展開した。(10)

筆者は最初、聴覚が音声中のデジタル成分を音素単位で抽出しているものと思っていたが、Liberman の指摘を受けて考えを改めた。そして、聞き手の脳は離散的に構築・発声された単語や句を、ひとつのアナログ刺激として、「音声的ではなく、音響的に」識別していることに思い至った。

3. デジタル符号語のアナログ処理

3.1 パブロフが前提にした犬の神経作用モデル

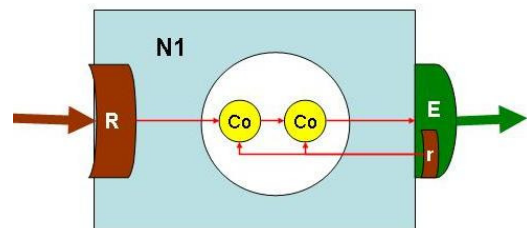
ロシアの神経生理学者パブロフ(I. P. Pavlov, 1849-1936)の名は、「条件反射」という言葉とともに世界的に知られている。日本では、彼が実験結果をもとにして行った講義録の邦訳が今でも文庫で購入できる。(11)

講義の第1講で述べているように、パブロフは「動物のこの器官(大脳半球)の生理学的研究は人間の主観世界の精密で科学的な分析の基礎となるはずである」という信念にもとづいて条件反射実験を行なった。

パブロフの実験の前提にはデカルトの反射概念がある。すなわち、「外界または生体内の要因は神経の興奮現象に変換され、興奮は神経線維を導線として中枢神経系に伝

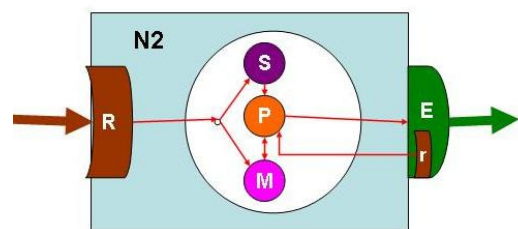
えられ、そこから以前に形成されている結合のおかげで他の導線に沿って作用器官へ運ばれ、そこで今度は器官の細胞の特殊な過程に変換される。さまざまな要因は生体のいろいろな活動と因果関係をもって法的に結びつく」というものである。外界の信号が神経を伝わって作用器官に伝えられるこの過程はたしかに反射である。

脳生理学者の時実利彦が脳内神経系のモデルをいくつか図示している。時実によれば、神経系とは、環境の状況や様子の変化を刺激として受け容れる受容器(感覚器, R)と、反応効果をおこす効果器(E, 筋肉や分泌腺)の間であって適切な反応効果のみちびく働きを荷うものである。(12)



(図 1)

N1 型の神経系は、受容器(R)で受けとめた信号を、定められた仕組みで運動や分泌の指令に変換して効果器(E)へ伝える伝導器(Co, コンダクター)の役割をする。この反応効果は、刺激に拘束された紋切り型であって、骨格筋にみられる反射運動や、内臓器官にみられる調節作用や本能・情動を含む。そして効果器のなかにある感覚器(r)やそのほかの感覚器から、フィードバックされる情報によってホメオスタシスが保障されるように、伝導器の働き方が調整される。(図 1)



(図 2)

本書を何度か読んでやっと理解できたことであるが、パブロフは一貫して N1 モデ

ルを前提にして実験し解析を行なった。一方、時実は、受容器で受けとめた信号を感覚(S)し、記憶(M)し、記憶に照らして感覚情報を処理し、運動や分泌の指令を出す情報処理・運動発現器(P)をもつ N2 型の「適応行動」モデルも提示している。(図 2)

パブロフは、条件刺激によって反射を引き起こしたので「条件反射」と名づけたが、そこに犬の記憶と情報処理が介入したと思われるので N2 モデルの「学習」や「適応行動」として理解することが適切と思われる。(7)

3.2 パブロフの行なったさまざまな実験

パブロフは、犬に聴覚刺激(笛, メトロノーム, クラクション, 机を叩く, ベル, オルガンなど)や視覚刺激(回転物, 電灯・光, 肉粉, 正方形・円形・楕円形, 階調の異なる灰色, おもちゃ)や触覚刺激(背や足の皮膚への機械刺激・温熱刺激)や嗅覚刺激(樟脳, 酢酸アミルの匂い)を「条件刺激」として与え、次いで無条件刺激(餌や異物である酸)を口の中に入れることによって、条件刺激によって無条件刺激に対応する効果(消化・洗浄用の唾液分泌)を起こすように「条件づける」ことを実験の出発点におく。(条件刺激は視聴覚刺激も嗅覚・触覚刺激もすべて遠隔刺激であり、口内を直接刺激しない。犬が信号と餌・酸をどう結びつけるかの実験であるので、N2 モデルであると考えerほうが妥当であろう。)

パブロフはこのように条件づけられた犬に対して、

- 1) 条件刺激を与えるがそれに続けて無条件刺激を与えないと、急速に効果が薄れていく「消去」,
- 2) 条件刺激に伴って無条件刺激を与え続けていても、ゆっくりと効果が生じなくなる「抑制」,
- 3) 条件刺激でない新しい刺激を与えてから条件刺激を与えると新しい刺激に対しても効果が生まれる「連鎖」, 「二次条件反射」,
- 4) それまで一度も条件刺激として使ったことのない新しい刺激を与えたときに、条件刺激として確立している刺激を与えても効果が生まれぬ「条件抑制」,
- 5) 条件刺激と無関係な刺激が存在すると、一時的に効果が生まれなくなる「外抑制」,
- 6) 条件刺激と無条件刺激の間に一定の時間をおくことによって効果も遅れて生じる「遅延」,
- 7) 遅延している間に他の刺激を受け、効果が無条件刺激に先行して生じる「脱抑制」,
- 8) 無条件刺激を伴う条件刺激と伴わない条件刺激を交互に与え、条件刺激のわずかな違い(色の濃淡や楕円率や周波数など)が効果の有無に反映される「分化」,
- 9) 分化によってつくられた無条件刺激を伴わない条件刺激のすぐあとで、無条件刺激を伴う条件刺激を与えると効果が増す「(正の)相互誘導」,
- 10) 無条件刺激を伴う条件刺激のすぐ後で、分化によってつくられた無条件刺激を

伴わない(負の)条件刺激を無条件刺激とともに与えても効果(唾液)が生まれず「(負の)相互誘導」などの実験を行なった。「」内は実験結果をもとに現象につけた名前である。

パブロフは正直な学者であり、すべての実験結果の理由を説明しきれたわけではないことが述べられている。その正直さによってパブロフの実験は現代的に再解釈される余地をもつといえる。

犬が適応行動・学習したのなら、たとえば上の 1)は、信号はもはや意味をもたなくなったと犬が判断したのであり、2)は犬が実験に疲れた・飽きた、と考えられそうだ。そしてパブロフが説明に困った 9)・10)の「相互誘導」は、犬の脳内記憶の中で、餌や異物を伴う信号と伴わない信号が、それまでの学習の結果、価値づけられて相互に関係づけられて体系を構築したと解しうる。つまり犬もアナログ信号を体系づけることを示唆していないだろうか。

3.3 パブロフの行なった複合刺激の実験

パブロフは第 8 講で「継時複合刺激(compound successive stimuli)」を論じている。これは「同じ現象が精密な形で生ずる」刺激である。

- 1) 同じ刺激の組み合わせを、異なるタイミング(リズム)で与える。
 - 1a) 同じ刺激(一定の音)からなる「音 1 秒、休み 2 秒、音 1 秒、休み 1 秒、音 1 秒」のセットが、5 秒の感覚で 3 回繰り返され、無条件刺激を伴った。
 - 1b) 同じ音刺激を、休みの位置を入れ替えて、「音 1 秒、休み 1 秒、音 1 秒、休み 2 秒、音 1 秒」として、5 秒の感覚で 3 回繰り返し無条件刺激を伴わない。
- 2) 同じ分析器(たとえば耳)に属する刺激が一定の順番で同じ長さと同じ休みを伴って配列されていたものを、順番を逆に与える。
 - 2a) ハ・ニ・ホ・ヘの音 で無条件刺激を伴う。
 - 2b) ヘ・ホ・ニ・ハの音 (2a の逆の順番)で無条件刺激を伴わない。
- 3) 別の分析器に属する 3 つか 4 つの刺激が同じ長さ同じ休止で構成されたものの順番を完全に反転して反対の順番にする。
 - 3a) 電燈をつける(L)、皮膚の機械刺激(C)、水泡音(S)で無条件刺激を伴う。
 - 3b) 水泡音(S)、皮膚の機械刺激(C)、電燈をつける(L)で無条件刺激を伴わない。
- 4) 別の分析器に属する 3 つか 4 つの刺激が同じ長さ同じ休止で構成されたものの順番を、そのなかの 2 つだけ位置を変える。
 - 4a) シューという音(S)、高い音(hT)、低い音(lT)、ベル(B)で無条件刺激を伴う。

4b) ベル(B)、低い音(lT)、高い音(hT)、シューという音(S)で無条件刺激を伴わない。

実験の最終成績として上のすべての場合に「はじめの複合と変化したものとの分化が形成され、変化したものは正の条件反射作用を失い、負の抑制性の刺激となった。」

しかし「これらの例のなかで抑制反射の形成はそのうちのあるものは非常におくれた。場合によって相対的な分化はかなり早く出現したが、完全な絶対的分化はひどい場合には数百回の反復を経てやっとえられた。」という。

とくに困難であったのは 4)の分化であるが、「課題を最終的に解決するためにはより簡単な複合刺激から段階を経て分化をつくることも必要であった。」だが、「完全な分化が形成されたあとでは、個々の成分を試みるとその正の条件反射作用を失っていた」というのは興味深い。

3.4 継時複合条件刺激実験の結果解釈

パブロフはあくまでも信号と大脳皮質細胞の現象と理解し、犬が記憶にもとづいて情報を処理しているとは考えないから、継時複合刺激の場合も「どのようにして同じ刺激の複合が、つまり大脳半球の同じ細胞に作用する刺激が慢性的に同じ期間これらの細胞に、あるときは興奮、あるときは抑制の過程というふうに異なった刺激となるのであろうか」と考える。

だが、犬も記憶・認識・判断すると考えるなら、もっと単純明快な解釈がえられる。犬は一連の継時複合刺激をひとつの符号として認識したのだ。刺激の順番が違っても、複合刺激をバラバラにして個別成分を開かせても、同じ符号として認識されないのだ。これはヒトの言語において、「いちご」と「ごちい」、「いごち」、「いち」、「ちご」が別の単語として認識されるのと同じだと考えればよいのではないか。

パブロフは「大脳半球皮質は、大規模なモザイク、言いかえれば非常に大きな信号板である」というが、継時複合刺激はそのモザイク状の信号板にひとつの軌跡パターンあるいはシンボルを描くのではないだろうか。

3.5 動物は概念はもつが、文法はもたない

パブロフのみならず動物の脳を使った実験は、ヒトと動物の大脳半球の基本的な構造や作用は同じであることを明らかにする。

3.3.8)で紹介したパブロフの分化実験の要領で、ウズラを使って、音素「t」と「d」を条件刺激とし、餌を無条件刺激とし、ついで回数を数えて、音素「t」と「d」を弁別できることを示した実験(13)や、チンチラに「t」と「d」を弁別させた実験もある。(14)それらは単に音を弁別したというだけでなく、条件刺激が意味と結びついて記憶されていることを示す。「概念は、さまざまな視覚刺激によっても形成されることが示されている。(略)人間の音声認識理論も、ウズラの行動を説明するのに求められ

るより以上に(以下でも)複雑である必要はない」(13)

動物も概念や文法をもつかということを実験した論文もみつけたが、「概念」・「文法」の定義がなく、「離散的組合せ」も定義がないうえに理解があやふやで、信頼できる結果は読み取れなかった。(15)(16)(17)「概念」と「文法」は、言語学や認知科学においても明確な定義がないまま、てんでバラバラに使われている。動物実験で「概念」や「文法」を確かめたいのであれば、きちんと定義しなければ、実験は成果を生み出しえないだろう。筆者は、動物の場合もヒトと同様に、「概念」とは「信号が記憶と結びついたもの」、「文法」とは「その信号そのものは記憶と結びつかず、単独では意味をなさないが、概念を接続・修飾・演算する法則を示す信号」と定義すればよいと思う。そしてアナログ通信である動物は、概念はもつが、文法はもたないと思う。

3.6 デジタル変調された信号をアナログ処理

さて、ヒトの音声言語認識においても、継時複合刺激は当然おこりうる。ならばデジタル原理にもとづいてユニークな音声刺激として構成された単語は、アナログなひとつのシンボル(あるいはパターン)として認識されているとは考えられないだろうか。単語は、音素単位でデジタルに復号化(復調)されて処理されるのではなく、ひとつのパターンとして単語単位でアナログ復調されて意味(=記憶)と結びつくと考えれば、音素処理能力の限界を超えた速度で話されても、単語の識別は可能である。

デジタルなものをアナログに処理するとは、たとえばオセロゲームや囲碁で石の数を数えることなく、盤上をひと目見て白と黒の石の差を言い当てるようなもの、コンサート会場や野球スタジアムの熱気で観客動員数を言いあてるようなものだ。どちらも簡単ではないが、経験を積んで熟練すればできるかもしれない。

脳で復調された単語は、単語と結びついた記憶(の総体)を呼びさすための見出しとして用いられるので、アナログ処理して1単語を1シンボルとして扱っても問題はなく、むしろそのほうが高速に処理できる。もちろん使い慣れない単語、耳慣れない単語は、丁寧に対応する必要がある。

4. 現代の聴覚生理学との整合性

4.1 聴覚メカニズムと大脳皮質一次聴覚野(AI)の周波数局在性

「聴覚メカニズムは、聴神経と一次聴覚野の間に4個から5個の神経核が存在していることと、対象となる刺激が時間変化を伴い複雑であるために、まだ十分に解明されていないところがある。」(16)しかしながら、これまでにわかった範囲では、パブロフが大脳皮質をモザイク状の信号板と考えたことは妥当であったと思われる。

外耳、中耳を経由して内耳に到達する機械的な入力刺激は、蝸牛において電気パル

スである神経刺激に変換されて出力される。「左右の耳に到達した音波は、聴覚末梢系で周波数ごとに分解され、位相固定されたインパルス信号に変換されて聴覚中枢系へと伝達される。」(18)聴神経は脳幹の蝸牛核で3つの蝸牛神経核に分岐し、そこからさらに分岐して、同側あるいは反対側のいくつかの神経核へと信号が伝えられる。

たとえば脳幹にある上オリーブ内側核(MSO)の細胞は、左右の両耳間の位相遅延に対応する。一方、「上オリーブ外側核(LMO)の神経細胞は、両耳における特徴周波数の振幅差を感知する」。これらの上オリーブ核の出力は、下丘中心核(ICC)と外側毛帯核(DNLL)に送られる。「DNLLの神経細胞の多くは、両耳から影響を受け、両耳間の時間と強度差に反応する。これに対して、外側毛帯腹側核(VNLL)への入力の多くは、反対側の耳の蝸牛核からもたらされる。それはVNLLの神経細胞の全部ではないが多くが片耳神経細胞から構成されていることに反映されている。VNLLの出力の多くは直接同じ側のICCに送られる。」(19)詳細な研究成果の一部を読むだけでも、聴覚信号処理の複雑さと研究の難しさが理解できる。

ICCの出力の多くは同じ側の内側膝状体(MGB)に送られ、最終的にMGBの出力が大脳皮質の聴覚野に送られる。ICCもMGBも一次聴覚野(AI)も、内耳有毛細胞の周波数局在性に対応して神経細胞が特徴周波数(CF)をもっていて、CFの値に沿って層状に並んでいて周波数局在性をもつ。

大脳皮質のAIの「皮質ニューロンは、刺激の過渡電流に対して、キビキビと、短時間、同期した反応を示す。しばしばそれらは申し分のない時間的精度と、著しいニューロン間の共時性をもつ。」パブロフが信号板と呼んだAIにおいて、音声の分散成分や振幅の包絡線に合わせて、神経細胞が発火しているのだろう。

4.2 中枢聴覚神経システムにおける話し声の処理

大脳皮質上で、どこまで音声言語処理が解明されているか、参考までに「耳の生理学」の中の一章「中枢聴覚神経システム入門」を読んでみた。

「周波数局在地図が、刺激の表現において正確に何を意味するのかについては議論の余地があるかもしれないが、聴覚刺激情報の中核『処理』が周波数にもとづいて行なわれていることは疑いがない。(略)ヒトの話し声に関する限り、皮質神経細胞が、声の絶対音程を設定した声門のパルスに同調した神経発火を復元することはありそうにない。しかしながら、話し声信号の音声的に重要な要素のタイミングを示すことができることは疑いようがない。この点において、最近明らかになったことは、話し声信号においてもっとも重要な時間的成分はよりゆっくりとした振幅の包絡線であって、正確な波形構造ではないということである。(略)話し声の皮質上での表現は、音声的というよりもむしろ音響的であり、声の絶対音程とも無縁である。先行する動物実験をもとにすればもっと多くのことがわかるであろう」(19)偶然にも筆者が思いついたこととほぼ同じことが専門家によって論じられているのは、心強いことである。

最新の聴覚研究の成果にも、あるいは動物の概念形成実験においても、パブロフの実験成果と実験手法は今なお意味をもっていることを感じる。

5. むすび

5.1 デジタル変調された音声をアナログなシンボルとして処理

聴覚は、デジタルに構成された単語を、アナログな振幅の包絡線情報にもとづいて、ひとつのシンボルとして処理している可能性について論じた。ヒトのデジタル言語システムにおいて、フォニットを重複順列で組み合わせて単語を形成するのは、離散的な単語数を増やすためであり、受信者は詳細なフォニット情報を必要とせず、単語を特定して、その単語と結びつく記憶を思い起こすことができればよい。デジタル変調された音声符号語を、ひとつのアナログ信号(シンボル)として復調し処理するのは、処理速度を速めることにつながるのだからメリットは大きいと考えられる。

デジタル情報を、デジタルに復調してそれを意味(記憶)へと変換するよりも、復調の段階でいきなりアナログ信号に復元してアナログな意味と結びつけることは、意味と信号の親和性もよくなるだろうから、実にスマートな情報処理手法である。

このスマートな手法は、人類が思いついたわけではない。人類はそれについて無自覚なのだから、同じことは遺伝子の情報システムで起きている。4種類の核酸塩基からなるメッセンジャー(m)RNAは、3つの塩基からなる64通りのコドンの組合せをトランスファー(t)RNAに伝えて20種類のアミノ酸への翻訳を促す。離散信号の組合せであるコドンは、翻訳に際して忠実にデジタル処理されているわけではない。なぜならば64通りのコドンに対応するtRNAの数は30から40種類だからだ。(1つのアミノ酸に対して2種類以上のtRNAが存在し、1つのtRNAは複数のコドンに対応しているため、30から40種のtRNAが1つの翻訳系で使われる。(wiki <http://ja.wikipedia.org/wiki/転移RNA> より))

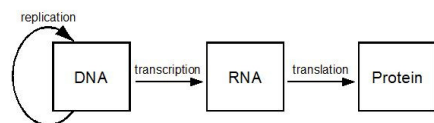


図 3

DNAからタンパク質への流れは、後戻りのない一方通行(セントラル・ドグマ)である。(図3) 言語の場合も、一般的通信モデル(図4)にならって情報源から到達先まで情報は一方通行に流れると考えることがふさわしい。単語を離散的な音声信号として大気中に送り出すまでは、繊細な発声器官運動制御が必要であるが、受信した後は、脳内の記憶を引っ張ってくるだけでよい。復調後に音素情報は必要とされないのだ。

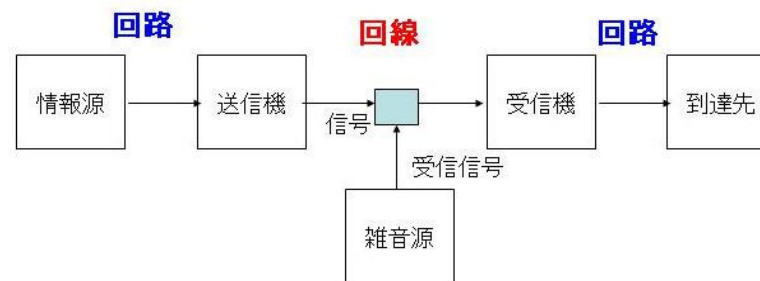


図 4

5.2 ヒトと動物の違いは音声通信のデジタル性のみか

筆者は深刻化する地球環境問題への関心から、人類の起源、文明の起源を求めて研究を開始し、結果的にそれはヒト言語のデジタル性の発見へと結びついた。今回の考察で、デジタル符号語を振幅の包絡線情報としてアナログ音響処理している可能性を見出して、ヒト言語のデジタル性の検討は、さらに一步進展した。

ヒトは、単語数を増やすために離散有限な音節を組み合わせるデジタル原理を用いているほか、概念を接続し修飾する文法法則を編み出して、長く複雑・精巧なメッセージを組み立てて順序だてて送るシリアル通信を行なうようになった。文法は、概念の演算子であり、五官の記憶と結びつかない抽象的意味をもつ符号語である。文法があるから、ヒトは概念の演算・操作が可能になり、科学的概念や歴史的な概念など自分の五官の感覚記憶にもとづかない抽象概念を獲得したのである。

遺伝情報は、mRNAがアミノ酸を指定し、リボソームがアミノ酸をポリペプチドにつないで、それが折り畳まれてタンパク質の三次元構造になる。そしてそれはさらに大きなシステムである細胞、器官、個体をつくりだす。片時としてやすむことなく続くこの現象は、生命の論理、生への衝動、生きる力によって突き動かされている。我々も、人間がもつ生命の論理装置に、もっとも人間にふさわしい論理を実装して、概念を文法によって接続し修飾し演算して、ポリ概念を生み出し、人類をより高い精神の次元へと導く思想、宇宙と生命の神話を紡ぎださなくてはならない。

ヒトも動物も、通信システムの機能や効率が違うだけで、ともに地球上で進化した生き物であることに変わりはない。ヒトが効率のよい通信システムを獲得したからといって、ヒトだけが偉くて、他の動物や植物の生命を好き勝手に殺りくし操作してよいということではない。ヒトは一日も早く、自分だけが偉いという誤った考えを卒業しなければならない。

レヴィ=ストロースは、ヒトが社会を築いて生きているのは、自らを犠牲にして集団を生かすための暴力装置「捕食」(食人)の文化が存在しているからだという。(20) 人間間の捕食を避けるために、動物を家畜化し、森林を切り拓き、他の動植物の命と生活の場を奪いつくしてきたのが人類の歴史であった。それはこれまでもエジプト、メソポタミア、インダスなどの地域文明の興隆と衰亡をもたらしたが、ついに地球規模での文明の衰亡が起きたのだ。

5.3 遺伝子情報システムを取り込んだ人類

ヒトが言語によって多くの過ちや勘違いを行なったからといって、言語を原罪として呪う必要はない。デジタル通信システムは、ヒトが創造したものではなく、先行する遺伝子の通信システムの模倣にすぎないのだから。

「遺伝子コードの構造と操作についての研究の方が、いくつかの単語あるいは文の断片をボノボに教え込むために注がれる労力(これによって猿についての知識を得ることにはなっても、言語について知ることにはならない)以上に、分節言語の性質について多くのことを明らかにしてくれるし、また逆もしかりである。

というのは、ヴィーゴの螺旋に似せて、同一の機能が、遺伝子と分節言語という生命体の異なった段階に回帰的に現われるからである。その獲得あるいはその学習の段階を遡っても分節言語の起源を発見することはできないだろう。分節言語のモデルとなる、予めそれを構成している別の言語がある。そしてその起源(さらに同じ属性を有しながらある水準から別の水準へ再出現するという神秘)が、心理学者あるいは言語学者だけの守備範囲にとどまらない諸問題を提起する。」(20)

20世紀の分子生物学や動物生態学の成果をもとに考えるならば、人類は安全で快適な洞窟の中で暮らすことによって、新生児が生後一年間ほぼ寝たまゝの状態で過ごせるようになり、おかげで他の大型霊長類より1リットルも大きな脳を獲得した。その大きな脳を使って、DNAをmRNA、そしてタンパク質へと結びつける遺伝子の通信システムによく似た言語を獲得した。地球環境問題が深刻化する21世紀に、それをどのように生かして、人間は人間らしく生きていくのかと考えなければならない。筆者の拙い考察が、人類がこれから前向きに生きていくための参考になれば幸いである。

参考文献

1) 得丸公明 ヒトの話し言葉はデジタル通信であり、情報源・通信路・伝送路の符号化・復号化

が行なわれている 信学技報 IEICE Tech Repo TL2009-28

- 2) 得丸公明 遺伝情報とヒト話し言葉の遺伝子型・表現型における意味づけの相似性について 情処学会 バイオ情報学研究報告 2009-MPS76 BIO19-48
- 3) 得丸公明 人類進化のデジタル符号化要求仮説：脳内にデジタル符号処理回路がつけられて人類文明は生まれた、情処学会音声言語研究会 SLP-79-41 2009
- 4) 得丸公明:人類言語はデジタル通信である、情処学会自然言語研究会報告 FI97NL195-5
- 5) 得丸公明:ヒト話し言葉デジタル通信システムにおいて情報源符号化をになう概念の形成・使用におけるパターン認識のはたす役割～条件反射の第二信号系=言語の生物学的構造～信学技報 TL2009-40
- 6) 得丸公明:ヒト話し言葉に適用されているデジタル原理について ～時系列上で次元配列されて送受信される離散・有限信号である音節、ならびにその情報源・通信路・搬送波(伝送路)における符号化処理過程について～信学技報 SP2009-169
- 7) 得丸公明:パブロフの条件反射実験の言語学的解析～脳内言語処理メカニズムについての試論～信学技報 LOIS2010-19 (2010年5月20-21日)
- 8) ウィーナー, N 「人間機械論 第二版 人間の人間的な利用」 鎮目・池原訳, みすず書房, 2007年, 原書 1950
- 9) Liberman, A.M, Cooper F.S., Shankweiler, D.P., Studdert-Kennedy, M, 'Perception of the Speech Code', Psychological Review, 74:6, pp 431- 461 Nov. 1967
- 10) Liberman, A.M. Mattingly, I.G. The Motor Theory of Speech Perception Revised. Cognition, 21 (1985) 1-36 この中で話者のジェスチャーが作用することが述べられている。
- 11) パブロフ 「大脳半球の働きについて」, 川村浩訳, 岩波文庫 1975
- 12) 時実利彦 人間であること, 岩波新書 1970
- 13) Kluender, K.R., Diehl, R.L., Killeen, P.R. Japanese Quail Can Learn Phonetic Categories, Science, 4 Sept. 1987, 237:1195-1197
- 14) Kuhl P.K., Miller J.D., Speech Perception by the Chinchilla: Voiced-Voiceless Distinction in Alveolar Plosive Consonants, Science, 3 Oct., 1975, 190:69-72
- 15) Kako, E. Elements of syntax in the systems of three language-trained animals, Animal Learning & Behavior, 1999, 27(1), 1-14
- 16) Pepperberg, I.M. Rethinking syntax: A commentary on E.Kako's "Elements of syntax in the system of three language-trained animals" Animal Learning & Behavior, 1999, 27(1), 15-17
- 17) Herman, L.M., Ueyama, R.K., The Dolphin's Grammatical Competency: Comments on Kako(1999), Animal Learning & Behavior, 1999, 27(1), 18-23
- 18) 伊藤一仁 初期聴覚中枢系の神経細胞レベルでの時間情報処理機構に関する研究, 2003: <http://hdl.handle.net/10119/934>
- 19) Phyllips, D.P. (2000) Introduction to the Central Auditory Nervous System, in A.F.Jahn and J.R. Santos-Sacchi (Eds), "Physiology of the Ear", 2nd Ed. San Diego, CA: Singular, pp613-638
- 20) レヴィ=ストロース, C. アメーバの譬え話, 出口顕訳, みすず, 2005.7