



Title	言葉と認識についての覚書(2)
Author(s)	浜口, 稔
Citation	明治大学教養論集, 216: 95-111
URL	http://hdl.handle.net/10291/8888
Rights	
Issue Date	1989-03-01
Text version	publisher
Type	Departmental Bulletin Paper
DOI	

<https://m-repo.lib.meiji.ac.jp/>

言語と認識についての覚書(2)*

——言語音の産出と理解のダイナミズム——

浜 口 稔

1. 序

言語と認識という概念を軸にして、コミュニケーション、すなわちヒトとヒト、ヒトと環境世界とのあいだのやりとりをベースにして、入り組んだ緊密な場として成立する一種の「コトバの生態」をイメージしていくような方向で、言語生物たるヒトの生存のあり方について概観してみようと思う。ヒトはコトバを媒介にして、心の組織化を不断に実践し、心のシステムを更新し、さらに再編していく存在である。今回は認識主体としてのヒトというイメージで、特に言語音の産出と理解の実際面に論述の焦点をしばっていく。

人工知能の開発であるとか、通信ネットワークの整備であるとか、さまざまな先端科学からの刺激を受けて、「コミュニケーションとは何か」という古典的な問題は、新たな装いをもって再考される技術的な素地が整いつつある。しかしながら、単純にコトバの産出と理解の仕組みというが、人工知能に自然言語を理解させたり話させたりすることが実際に困難をきわめ、それをなんとか実現させようと相当な数の研究者が世界的なスケールでその試みに携わっている事実をふりかえてみても、われわれの日常感覚ではなんの変哲もない自然な過程であるコトバというものが、いかに複雑精妙なからくりのうえに成立する事実であるか、それだけでも思いをあらたにできる。とはいうものの、われわれの分析的な知性はかなりの程度コンピュータによるシミュレーションを期待できる状態にまで、あるいは少なくともそのような見込みのもとで人間精神

のもろもろの相を数値处理的にモデル化しようという気運が高まるほどに、技術の繊細さは進展を見せていることもまた背景にある（もっとも、皮肉な言い方をすれば、その技術に合わせて、われわれの分析的な知性の組織化のあり方を調整したというべきかもしれない）。ようするに「技術を媒介のモデル」が、以前に比べて非常に立てやすくなってきているということが、大変に大きいのである。伝統的な思弁の正当性が、とりあえずはコンピュータという記号学的・言語学的理念の権化を手段にテストされようとしている、と言い替えてもよいかもしれないのだ。

さて、ヒト同士の情報のやりとりというのは、ここでは単純に日常会話のやりとりを思い浮かべていただければよいのだが、コトバを発する話し手も、それに耳を傾ける聞き手も、実際には、言い淀み、言い間違い、表現の途切れ、などなど、断片的で気紛れな不確定要素にさらされ、さらには環境世界のおびただしい非言語学的ノイズに溢れかえる混沌の海にどっぷりつかっている。そのような中から、どうしてこうも首尾よく理解につながる言語秩序だけを拾い上げ、結果的にはコミュニケーションを成立させるに至るかを明確にするのは、実に想像を絶する厄介な作業であるといつてよい。まずは言葉を話し理解することの、もっとも物質的・生理的な基盤にある言語音の産出と理解について検討してみよう。

2. 言語音の産出と理解の仕組み

いうまでもなく、言語音を発する際には、普通は食べたり飲み込んだりするための「舌」や「唇」や「歯」、また呼吸のための「肺」や「気管支」などを利用する。これらはみな他の動物にも普通に備わっているから、そういうわけで言語音の産出は、もともとあった器官を利用していることになる。ただし形や機能において、その柔軟性やコントロールの容易さは、他の動物と比べて比較にならないくらい発達しているようだ。古典的な例であるが、以前、アメリカの心理学者のヘイズ夫妻が、生まれたばかりのヴィッキイという名のチンパンジーを引き取って育て、人間の言語音を発声させようとして大変な苦勞の末に、

ようやく、*papa, mama, cup, up* の四語を発することができるようになった程度であったという報告がある¹⁾。ヒトのさまざまな発声器官の筋肉活動の自在さ、さらに脳からそういう筋肉活動へと連動する神経解剖学的な条件は、ヒトと他の生物を隔てる決定的な差異のひとつである²⁾（音の発声と理解の仕組みが入り組んだ神経解剖学的な背景をもっている事実については、あとで興味深い観察例を挙げて考察しよう）。あるいはもう少しまとめるなら、声帯を振動させ、口の中の音響共鳴器たる口腔やら咽頭腔やらの空洞を実に柔軟に変化させ、また舌や唇や歯や顎を巧みに操ることで、音の値を定め、結果的にはいろいろどりの母音や子音を生み出している事実、この複雑微妙な諸機能の総合的な運用が、ヒトと他の類人猿とを分かち決定的な違いである、といてよいだろう。この総合の力が、人間の言語能力を地球上でユニークなものにしているわけである。ただ器官があるというだけではだめである。その器官が脳の言語中枢と密接に結び付いていなくてはならない。これが神経解剖学的に（神経パルスの発火の順序、神経の伝導の状態などの点で）組織化されていなくてはならないのだ。なにはともあれ、ここでは言語音を発することが、神経解剖学的に考えても、いかに微妙かつ高速な処理過程であるか、とりあえずはそのことを念頭に入れていただきたい。その事実を、発声器官、聴覚器官の解剖学的な精巧さと併せて、これから述べる言語音についてのいくつかの観察実験例でいっそう鮮明なものになると思う。

言語「音」は、端的に言って空気の振動現象であるから、当然ながら肺からの呼気を主に利用する。呼気というのは、通常は、気が付かないくらい静かなものであるが、音をそれで合成するとなると、まずは肺からの空気圧をうまく調節し、呼気の通路にある途中の声帯の振動を促さなくてはならない。この声帯が一種のゲートの役割を果たして、結果的にはさまざまな言語音の源となっている。言ってみれば、ここに、世界を言語的に体制化（分節化）する源のひとつが存在するわけである。

こうしてわれわれは主として、さまざまな発声器官を動かして、結局はわれわれの言語活動を微妙多彩な形の音でいどっていくのだが、この音の産出の

仕組みは普通に考えるほど明快なものではない。ヒトの発声器官というのは、当然ながら、それを作っている素材が柔らかい肉質のものであるから、そういう意味でハードな機械のように、精確には、調音の際の「音価」を徹底できない。おまけに、音が口を離れて空気中に放り出されると、物理的にはただの空気の振動となってしまう。それに加えて、こんなことを考えに入れていただきたい。つまり、口から放出される言語音は、もともとは脳で組み立てられた伝達内容という「精神的」合成物であったわけであるから、これが空気の振動であるにすぎない音の中に、どのような形で織り込まれているか、ということである。別な言い方をするなら、当初の心理的構築物が電磁化学的な過程へと変換され、さまざまな発声器官を動員し活性化させ、空気振動という純物理的現象へと変換させたいうで改めて総合化される際に、その当のメッセージは、聞き手の理解を損なわない程度にその内容を保持したまま空気中へと放出され、最後には耳の器官へと伝わっていけるだろうか。それをわれわれの聴覚器官は、

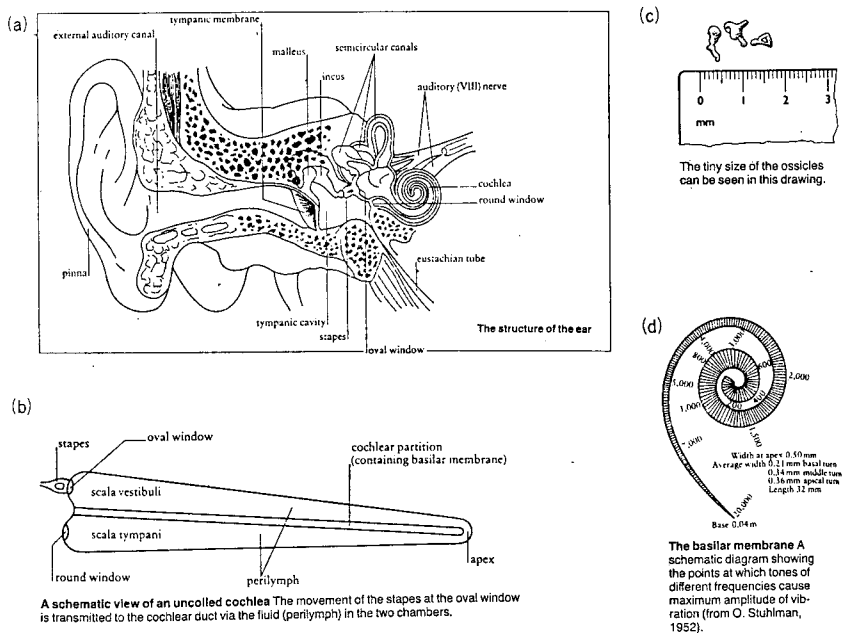


図1. 聴覚器官の仕組みの数々

おびたしい雑音が飛び交う外界から、言語音の振動だけを聞き取らなくてはならないわけであるから、そのからくりはそれだけを考えに入れても大変に不思議である。しかしそのような不思議は、耳に届いたあとの解剖学的な経路をさらに概観すれば、なおいっそう際立ったものになるのだ。

聴覚器官というのは、一言でいうと、「外耳」「中耳」「内耳」のそれぞれの器官で外から入ってくる音響パターンを順次処理し、最終的には神経電氣的パターンに変換する一種の精巧な「変換器」である（この点は発声器官の大ざっぱなミラーイメージをなしている）。つまり聴覚器官の物理的機能は、外から来る音響的な振動を受け取って、耳内部の力学的な機構を經由して聴覚神経に至らしめ、それを脳の言語中枢につなぎ、最終的には音として認知させるための担体としてのものである。まずは前頁の図1をご覧ください³⁾。

さて、外から集められた音波が、まずは鼓膜に当たるとというのが、音の知覚のそもそもの発端になる。中耳は周知のように、三つの小さな骨（malleus「槌骨」、incus「砧骨」、stapes「鐙骨」）から成り立っている。この三つは中耳内部で中吊りになっていて、鼓膜の振動を巧妙に調節して「前底窓」(oval window) と呼ばれる膜に伝える役目を果たしている（この三つの骨は一種の「てこ」の原理で、鼓膜の振動を増幅させる働きをもっているが、その増幅はおよそ35倍にも及び、そのせいでどんな微かな音でも聞き分けられることを可能にしているようである）。その「前底窓」を叩く連続的な圧力刺激が、図にある蝸牛 (cochlea) の中の液を動かして聴覚神経につなぎ役割を果たすのだ。この液がブルブル震えて、この蝸牛の内膜に生えている毛状の神経細胞が「そよぐ」。この毛状の細胞は聴覚神経に接続していて、その神経が「そよぎ」の結果生じた「電気刺激」を脳に伝えるわけである。

ここでもう一度、この聴取の力学的・神経学的な過程を考えてみよう。まず外耳で集音された振動音は、もともとは特別な言語的なチャンネルがあるわけではない空気で満たされ開かれた空間に存在していたものだ。したがって、そのような空間を伝わっていく非言語的ノイズを始めとする音の複合体が、鼓膜に届くことになる。それを前庭窓に伝える内耳の三つの骨の力学的な機構は、

一見すると至極単純である（骨をつくる伝達素材としての性質のほどは、とりあえず考慮に入れない）。それが内耳の中の液状物質を微妙に振動することで、言語音を初めとする音楽、車の音、工事現場の騒音、知人の声などの、環境世界のおびただしい音を聞き分けることができるのであるから、それを担う神経解剖学的な仕掛けは、当然ながら、これまで概観した以上に精緻巧妙な構造をもつことを、絶えず念頭に入れていなくてはならない。しかも、こうして次々に入ってくる圧力刺激パターンを基底層で受け止めて、周波数だとか強度であるとかの聴覚体験を合成するために、脳はリアルタイムに解析していなくてはならないのだ。そしてその解析したパターンを、聴覚神経を經由して脳へと送り込み、こうした情報の聴取を脳は非常に速度と高い精度で達成しているのである。耳は、つまりはわれわれは、こうして聴覚神経による処理を経て、基本的には音の「高さ」「大きさ」「質」「長さ」という四つの心理的次元を統合して「聞く」という聴覚の過程を日々体験（＝解釈）していく。

音の産出に話を戻そう。「声帯」は、非常に広い幅の周波数帯をもつ振動のパターンを、強度や大きさを含めて一度に生成する。もう一度繰り返すが、われわれは顎や舌や軟口蓋や歯などを運動させ、共鳴箱としての口腔の形状をいろいろに変容させることができる。そうすることで「声帯」で作られたかなりの幅の周波数帯をもつ振動パターンから、リアルタイムに、さまざまに変化する形状に合わせて、そのつどの周波数帯が連続して共鳴し増幅され、一定の音質をもった言語音として、続々と外界へと放出されるに至る。

さて母音というのは、典型的には、口腔の形状変容を主たる調音活動にして形成されている。ようするに、口腔を一種の共鳴箱にして、声帯で生成される幅広い帯域をもつ音源から、形状に合う帯域を選んで抽出をできるように筋肉活動を制御し、実際の音として合成している。したがって、どこからどこまでがどの母音で、どこからどこまでがどの母音というぐあいに、しかも口の中の形をそうやすやすと急激に変化できるものではないという筋肉活動上の制約もあって、絶対的に定まった形がない、つまりは絶対的に安定した数値をもつ周波数の音というものがないことが、当然ながら予想されるわけである。機械と

違つて、人間の言語発声は、実際には、相当に輪郭のフェジーな過程である。非言語的な混沌の中をこのようなフェジーな音連鎖が伝わっているわけであるから、これも、われわれが常識的にイメージする「聞く」という所作の実現を奇跡的なものに見せてくれる驚異的な事実といえるだろう。このことは、特に図示はしないが、スペクトログラフで記録したスペクトログラム上のフォルマント分布を見ればよくわかる。

そのような記録を見れば明らかであるが、子音というのは、舌や歯や唇などを口腔内のいずれかの場所に接触させたり近接させたりすることでその存在が認知されるくらいのものであるから、その姿をきわだつた形ではしるしてはくれない。せいぜい母音のフォルマント構造の端の方で、わずかにまとわりつくような形で存在する。つまり、われわれの知覚が捉えられるくらいに十分な時間を占める子音はない、ということである。早い話、子音は母音なしでは知覚されないのだ。

ころろみに、*ba* という音のつらなり（正確な言い方ではない）をテープに採取して、そのうちの *a* の音だけを切り離すつもりで、*b* 音の側から慎重にハサミを入れて順次音をつめていく。そして順次これを再生してみる。そうすると、全体の音の *b* の音は次第に消えていって *a* の音しか聞こえなくなるのが観察される。これは予想どおりである。今度は *a* の側から段々に切っていく。ところがそうやっても、*b* の音は決して聞こえてくるようにはならない。言語音とは似ても似つかない音が、ちょうどフォルマントの推移点あたりに相当する部分で聞こえるだけのようなのである⁴⁾。つまり、通常われわれが「子音」と呼んで独立の実体であるかのように感じている音素なるものは、音響的には存在していなかったことになるのである。

もうひとつ、われわれの心の構えが、言語音の錯認を生む別な例を挙げておこう⁵⁾。*slit* という音の連続をテープに記録する。そして *sl* にハサミを入れて、そこに何も記録していないテープを足して隙間を作る。そうするとわれわれは、その何もないはずの隙間に音を聞き取って（あるいは捏造して）、*l* の音を補って *split* と聞いてしまうのだ。さらに *b* だけではなくて、*sore* の場合は、*t* を

補って、*store* と聞いてしまうというぐあいに、環境によって補われる音が違うらしい事実も確認されている。この事実からいっても、言語「音」の認識は、われわれの心の「構え」にたつぷりと汚染されていることは明らかだろう。

それにしても、音響的には何も存在しない隙間にある音を聞いてしまうという事実は、消極的な意味での「錯誤」である以上に、先ほどから強調してきた劣悪な言語環境をかいくぐってコミュニケーションを成立させる話し手・聞き手の能動的な「虚構の力」、別な言い方をすれば、確率や習慣を通じて培われてきた蓋然的な認知データをよりどころにした瞬時的かつ連続的な仮説の適用をして、脳がこの錯誤現象を積極的に利用していることの証拠として捉えても差し支えないように思う。それだからこそ、われわれのコミュニケーション活動は、一種の検証活動を連続的に交わすことで、聞き手・話し手のあいだに張りつめられる虚の（仮説の）空間を創造する、きわめて演技的な自己投機となりうるのである。

3. 言語音の認識過程

これまではもっぱら、単一の音の認知だけに話を限ってきたが、勿論現実には話される言語は、そんなふうには単発的に発せられるものではない。単一の音のつらなりという言い方からして、正確ではないのである。実際の会話というのは、言うまでもなく、さまざまな言語音の連続的な継起である。まずは単純なイメージになるが、ちょうど紐をとおしたビーズ玉のように、音の単位（つまりは音素）が次から次へと派生していく事態を想像していただきたい。あくまで便宜的にはあるが、われわれは少なくとも、ひとつの発話を伝統的にはそのように表記してきたのである。話し手はビーズ玉のひとつに一定の値をもつ音を封じ込めて、聞き手はそのひとつひとつの値を高速かつ正確に読み取って全体の（語の、文の、談話の）音形全体を連続的に把握するに至る。ところがそのようなイメージはまったく通用しないというのが、言語の産出と理解の実情なのである。このことは先のさまざまな観察・実験例からも明らかであった。これは自己充足的実体としての「音素」という概念へのきわだった反証例

となっているのだ。言い替えるなら、言語音は絶対的な値を持たず、随分使い古された言い方であるが、それが生起する環境にほぼ全面的に依存的なのである。

さてそこで、脳がリアルタイムで言語音を産出している過程を、また別の角度から検討してみよう。言語音はそれぞれ、その音を発するためには、違う器官内の神経や筋肉を利用することは先に述べた。脳からの音生成の指令は音によって違ったものになる。ようするに、言語音はそれぞれ、それにかかわる筋肉活動を統御する脳の中枢部での神経指令の組み方が全部違う。しかもそれぞれが、調音される場所に至るまでの経路の長さが違う。上記のビーズ玉方式で考えると、それぞれ音のビーズ玉をつくる神経指令も、神経解剖学的な経路も、その電氣的指令の伝導の速度も距離も、それぞれで異なっている。したがって、それぞれの音の発声の神経指令は、単純には、実際に口をついて出てくる音の順序に従って、次々に「発火」されるというふうに考えるわけにはいかないのである。簡単ではあるが、目ざましい例を挙げよう。下の図2を見ていただきたい⁶⁾。

たとえば、*pa* という音のつらなりを産出する事態を考えてみる。この図で見ると、*a* の音を出すための脳から喉頭の筋肉（図のc）に至るまでの神経経路の距離は、両唇音である *p* 音を出すための唇（図のb）や顎（図のd）までの距離に比べるとずっと長くて細い（つまり伝導効率が悪い）。そうすると、こ

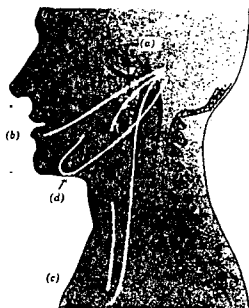


図2 Schematic cross section of vocal tract showing nerves connected to muscles. From Lenneberg 1967.

の音のつらなりを産出するためには、神經的指令の発火の順序を逆にしなくて
はならないことになるだらう。このような音のつらなりが大量かつ長時間生み
出される状況を考えてみるとよい。その場合話し手は、このおびただしい数の
神経解剖学的な経路を通して伝わっていく電気の経路や速度を考慮に入れつつ
(勿論無意識にはあるが)、神経的指令の発火の順序を猛烈な速度で調整し演
算実行しなくてはならないのである。もしこのような演算をそのつど改めてや
り直すというのであれば、脳の神経生理的な負担は途方もないものになるだろ
う。脳は疲れを知らないシリコン・ロボットではないのである。

これでヒトのコトバ活動が、神経生理面に話を限っても、いかに驚異的な事
実であるか了解いだけたことと思うが、聴取の速度について、もうひとつこ
ういう事実を検討してみよう。実際のヒトの言語音の解析度は、一秒につき大
体20から30単位であるらしい⁷⁾。たとえば、*bat* という音連鎖の認識は、そうい
う速度でなされている。しかし一方で、ただの音の連鎖の分析の力となると、
一秒ごとに7から9単位らしいから、これを埋め合わせるなんらかの認知の上
での戦略があることが考えられる。考えられる戦略は、20から30の単位を適
当な数ずつまとめて、大まかに7から9の音節に区切っているというものであ
る。たとえば、*bat* はひとつの音節単位であるかのように、*bat* という単位
を、聞いたときにリアルタイムで、いちいち*b*、*a*、*t* というぐあいに分析しな
いで、ワンセットまとめて脳に登録していて、そのようにして、結果的には、
最大20から30単位までの認識を可能にしている、というような説明である。早
い話、意味のないでたらめな音のつらなりでは、そのような効率のよい認知は
望めないことは、常識的にも予想できる。言い替えるなら、現実の発話は逐次
的認識によるものではなくて、既得知識をベースにした連続的な仮説の適用か
ら事前に音イメージを形成しておくような予見的・構成的認識である。平穩無
事なコミュニケーションが成立している限りは、習慣による反復的認識なので
ある。かくしてその限りにおいて、論点を先取りするなら、正真の意味でのコミ
ュニケーションは成立していないことになるのだ(つまり聞き手に届くメッセ
ージは真の「外部」からやって来たものではない)。

もっと別なレベルにおいても、たとえば、子音というものが母音に寄生して初めて音として聞こえてくるという事実を考慮に入れれば、コミュニケーションの成立はやはり錯認による一種の虚構によるものである。さらにまた、*bat* という音の連鎖がひとまとまり脳に登録されているのであれば、ピーズ玉のような音の連鎖という見方自体にも問題があることは明らかだろう。つまりこれも通俗的に言われていることであるが、*bat* を構成している *b*, *a*, *t* は、音響的には存在しない言語学者の抽象（虚構）である。繰り返しになるが、聞き手は、音の要素を聞き分けるに当たっての手がかりを、脳の中にワンセット持っていて、聞こえてくる音波の連鎖を、そのセットの手がかりと照合しながら、連続的に音を予測していくのだ。さらに言葉を替えるなら、聞き手は、音の要素を聞き分けるにあたっての手がかり、つまりはその言語音が発生する環境を脳内での音群のセットと、聞き手が置かれている場面から届いてくる音の振動のぐあいとを一体化するような形で「システム化」していて、それが言語動物たるヒトの交流のあり方として表象されている中で、リアルタイムで音を解読していくヒトの言語行動のダイナミズムを決定しているのである。ようするに、脳内表象を外界からの刺激に照合させて検証することで、外界認識が織り成されていく（自己組織化されていく）、と考えられるのである。

さてふたたび、音の単位というものが音響的には一貫した存在ではないことを確認できる、さらに追い打ちをかけるような、このような事実を見てみよう。もし子どもの言語の知識が、他者を媒介して、ようするに大人から、典型的には母親から「伝わってくる」というのであれば、その伝達のチャンネルはどのようなものとなるのだろうか。それは変動のない信頼に足る「経路」ではないことは明らかである。しかも音を発する「話者」の発音器官は恒常的な安定度を確保できるものでもない。ようするに、先に述べたような超高性能の言語器官を用いてもおぼつかぬくらい、実際にはわれわれの日常の場面を飛び交う言語音は不安定であることが、さまざまな実験報告といちいち照合しなくても明らかなのである。これは簡単に予測がついたはずのものなのであるが、ようするに、コミュニケーションがヒトとヒトとのやりとりのうちに成立するも

のであるからには、発声器官の個人差というものを考えなくてはならないということなのである⁹⁾。発声器官が違えば、つまり大きさや形が個人個人で違えば、当然結果として出てくる音の音響的な特質が個人個人で違うことになる。しかも大人と子どもではさらに違う。女と男も違う。声帯の大きさが違うから音源自体も違う。年寄りと子どもでも違う。おまけに子どもは発声器官のサイズをどんどん成長変化させていく。もしわれわれの耳が言語音の微妙な音響的な差異を聞き分ける精密機械であるなら、このような個人間の違いを実質的にはどう克服しているのだろうか。逆説めいてくるが、聴覚構造の精緻な聴取の仕組みは、厳密であればあるほど、言語音の認知の大きな障害になるとも言えるのだ。子どもの脳がもしタブラ・ラサであれば（これはありえないことであるが）、母親と子どものあいだに立ちはだかる「言語的混沌」たるさまざまなノイズに溢れ返っている音響のカオスを收拾するのは不可能であろうし、教育プログラムに則っているわけでもない母親からの情報は、見方によっては雑音以外の何ものでもないし、声帯や口腔の大きさの違いからくる音質パターンの相違はいかんともしがたいし、成長していく子どもの日々の発音器官の変化は止めようがないわけで、そのほかさまざまな不安定要因を考慮に入れると、コミュニケーションは実際には成立しているはずはないとも思われてくるのである。

結論を言うと、われわれが意識的に知っている他者との純然たる意思の疎通は事実上存在していないのである。われわれは常識的な意味での他者とリアルタイムで実質的な情報のやりとりをしているわけではない。早い話われわれは、「外部から」の情報を「内部で」捏造しているのだ。外界のなんらかの刺激を手がかりに、内部で連続的に仮説を立て、コミュニケーションを「虚構」している。そしてこの一種の生態力学的な虚構をディスクールの空間にして、私的かつ公的に言語生物として生きているのである。

これに関連してまた、こんな実験がある⁹⁾。まず *hit* という音を聞き手に向けて発してみる。ただしその際に、その音を聞き分けてもらう前に、いろいろな人の声で *Please say what this word is* という前置きを聞かせる。そうすると実験を受けた人は、その前置きの声によって、*hit* を違った音に聞いた

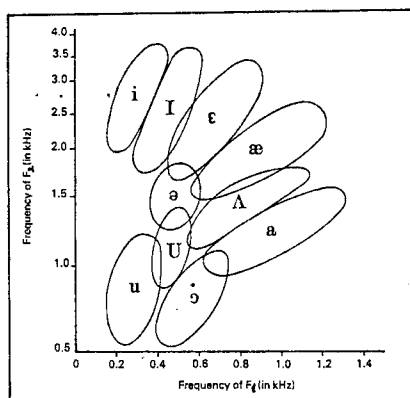


図3 Diagram showing the results of Peterson and Barney's (1952) study of formant frequencies F_1 for vowels spoken by 76 different speakers.) Note that the F_1 axis is a logarithmic scale 1kHz=1000Hz.)

らしい (たとえば, *het* と聞いたりしたらしいのだ)。そこで次頁の図3を参照してほしい¹⁰⁾。

これは76人(男33人, 女28, 子ども15人)に英語の音を記録させて, 個人個人のサウンド・スペクトログラムをまとめたものであるが, 個人差があることは明らかであり, 英語の言語音の絶対的な「値」というものが存在しないことは, どうやら間違いない。それならば, われわれはどうやって, 違うヒト同士の言葉のやりとりを可能にしているか, ということがなおいっそう不思議なことになるのである (この場合も耳が精密機械であることが前提とされる)。言葉を換えるなら, 物理的には異なっているにもかかわらず, 大人の *a* も子どもの *a* も同じ *a* であることを, 何を手がかりにして認識しているのだろうか。

この疑問に答えるひとつとして, こんな説明がある。古くから, 母音の三角形といって, *a*, *i*, *u* という安定した音があり, これは世界のたいていの言語に含まれる音らしい。たとえ三つしか母音がなくとも, その母音は少々ずれていても, *a*, *i*, *u* という基本三母音は保持しているのだそうである。そして図を見れば分かるように, 確かに, ちょうど母音の全体の分布の三つの端を, 当該の三母音が全体を囲むように位置づけられているのが確認できる。だから,

その三つを軸にして、話し手の母音全体のパターンを読んでいるのではないか、ようするに、ちょっとその人の話すを聞けば、はじめにその人の三母音の音の値を定めれば、残り全体の音のパターンは一挙に計算して出せるのではないか、という仮定を立ててみるわけである。もしそうであるなら、たとえ相手が子どもだろうが、大人だろうが、イントロの声を少しばかり聞けば、たとえば先ほどの *Please say what this word is* という声を聞くだけで、そのヒト固有の言語音のパターンの全貌を一挙に掴んで、つまりは初めて聞く人の言語音のシステム全体を瞬時的に見抜いて、結果としていろいろなサイズのヒトとのコミュニケーションを可能にしているのではないだろうか、というわけなのである。もしそうであるなら、いずれにせよこれは、われわれが暗黙のうちに、ひとりひとり関係論的な音のシステム固有の均衡状態を張りつめさせ、あるいは他者の（音の）システムを察知していることの有力な証拠であろう。

4. 展 望

これまで、言語の理解と産出の神経生理・解剖学的な側面、そしてその知覚のうえでの問題に議論を集中させてきた。勿論これだけでは、言語音の産出と理解をすみやかに過不足なく遂行できる事実を説明し得ているわけではないことは明らかである。これまで少しほのめかしてきたように、言語の理解と産出を音のレベルだけで考えるから、そういう問題が生じるのだ。*bat* の音の知覚のところで論じたように、この単語を不可分の統一体として記憶していれば、それを発するときも、聞き分けるときも、音響的には単一の音のまとまりとして聴取され、心理的にはこの単語を構成する三つの音は実在することになる。ようするに、もっと大きな、単語のレベル、文のレベルをも考慮に入れば、われわれの発音器官の性能にまるっきり寄り掛からなくても、コトバを理解する助けとなるものは、他にも見いだせるということなのである。コミュニケーション成立のための諸条件を必要なだけ揃えるだけでも、当然予想されるように膨大なものになる。だからこそ、それをまとめて処理するコンピュータの発達と人工知能の開発は、この膨大な条件を一気に実行させて、人間の

コミュニケーションの全貌を展望する最強のツールとして注目されているわけである。コミュニケーションの問題が新たな装いのもとに模索されつつあるようになってきているのは、以上のような事情が背景にあるとあってよい。

ともあれ、単純な音の理解ばかりでなく、ちゃんとした伝達内容をもつ言語音を大量かつ高速度で理解しようとなると、これまで見てきたように、解剖学的、神経学的、音響学的にかかわっている部分を覗き見るだけでも、緻密繊細な局部とそれが相互に連携する途方もない学際的な広がり进行を思い描くことができるだろう。ヒトはヒトとヒトのあいだの物理的（非言語学的）空間に「音」を放って意思を伝える。繰り返しになるが、そのすきまはおびただしいノイズであふれ返っている。コンピュータ同士のコミュニケーションなら回路をとってノイズに遮られる気遣いはない。しかし回線を介さないヒトとのコミュニケーションということにもなれば、やはり音声認識をはじめとするさまざまな言語研究の成果を取り込んで、認知の仕組みと連動させたセンサー技術の開発を進める必要が出てくるだろう。それに加えて、個別言語についての知識、つまり言語音、単語、文法、さらには常識、作法、文化、習俗、社会もしくは世界についての知識などを当然総動員しなくてはならないわけで、そういう意味で、コトバ活動のカオスから抽出されうる、おそらく行動する生物としての知能構造に由来する複雑な動的システムの一環としてヒトをイメージし、その範囲で言語音の産出と理解を捉え直さなくてはならないのである。かくしてコトバの研究は、絶え間なく運動する生のプロセスとして動員される驚異的な複合体である人間の存在様式の全容を科学する材料を集結させなくてはならないがゆえに、形として幅広い学際的な連携が必要になるのである。それでなくては、コミュニケーションという実に豊かな「ふくらみ」をもつ現象を把握できるようにはならないとあってよい。

ようするに、われわれの認知の仕組みは、コトバを媒介にして外の世界から知識を得るという事実にかかわっている。知覚の問題、認識の問題、思考の問題、判断力の問題、などなど、太古の昔から、われわれが生きて、外世界から知識を摂取し、蓄積して経験を深め、世界をある一定の様相のもとに「知る」

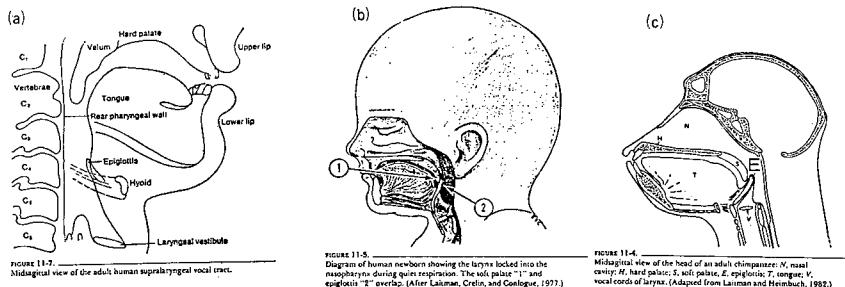
よくなることの不思議そのものにかかわるのである。言語活動であふれ返った日常の情報空間が、われわれの知覚を媒介にして、どのように体制化されているか、いかにそれは生物としての性能と適応能力のうえに成立しているか、コトバと認識をめぐる考察を諸科学との関連のもとで、できるだけ幅広く総合的に展望することの意味はそこにある。

〔註〕

* この論稿は、*Meropolitan* vol. 27 に掲載された拙著「言語と認識についての覚え書」の続編である。

- 1) Hayes, C., *The Ape in Our House* (Harper & Row, New York, 1951) を参照。動物のコミュニケーションに関しては、そのあとでもかなり大がかりな実験プロジェクトがあって、われわれのヒトの言語を教えようという試みには実に長い歴史がある。動物のコミュニケーション能力を検討することで、同じ種類の生物同士が何らかの情報を交換する事態とはいかなる事態であるか。あるいはまた、人間と動物のコミュニケーションはどの程度可能であるのか。そのようなコミュニケーションに比較してヒトの場合はどう位置づけられるか、そのように設問を立てて、人間の言語に関して何らかの展望をもとうという研究プログラムは少なくない。チンパンジーには言語音を操るための解剖学的・生理的基盤が十分整っていないことは確認済みであるから、当然ながらチンパンジー同士のコミュニケーションにおいてメッセージを乗せる容器が特に「音」でなくてはならない必然性はない。実際に「耳の不自由な人々」用に開発された「手」と「表情」を巧妙に組み合わせた「アメリカン・サインランゲージ」(ASL) と呼ばれている人工言語を使った実験・観察例があり、かなりの成果を挙げている。主な資料としては、Allan M. Schrier et al., *Behavior of Nonhuman Primates, Vol. 4* (Academic Press, 1971). Eugene Linen, *Apes, Men and Language* (Penguin Books, 1974), Rumbaugh, Duane M., ed., *Language Learning by a Chimpanzee: The Lana Project* (Academic Press, 1977), Herbert S. Terrace, *Nim: A Chimpanzee Who Learned Sign Language* (Washington Square Press, 1979). F. Patterson and E. Linden, *The Education of Koko* (Russell and Volkening, 1981) などがある。ただし、Terrace の場合は、チンパンジーの Nim が言語を習得したことを否定する発言をしているようだ。Thomas A. Sebeok and Jean Umiker-Sebeok, eds., *Speaking of Apes: A Critical Anthology of Two-Way Communication with Man* (Plenum Press, 1980) は、類人猿の言語実験に批判的な論文を集めたアンソロジーである。
- 2) もし言語音を発声するための生理解剖学的な基盤が充分ではないと言うのであれば、発声器官へと張り巡らされる神経解剖学的連絡網の欠如ということを別にしても、さらにはノーム・チョムスキーがいうような「種(ニヒト)に特有の」生得の

言語能力が存在しないことを別にしても、言語音の調音をスムーズにコントロールできない幼児をはじめとするサル、イヌなどが、「話す」ことができない事実の説明のひとつにはなる。たとえば、以下の図に挙げる咽頭腔 (pharyngeal cavity) が十分に存在するのは、成人したヒトだけである。



(a)のヒトの大人の Epiglottis は、それぞれ(b)のヒトの幼児の②, (c)のチンパンジーのEに相当するが、後二者の Epiglottis は軟口蓋に近接している。しかし(a)では距離があり、そのあいだに咽頭腔が発達していることに注意していただきたい。Philip Lieberman, *The Biology and Evolution of Language* (Harvard Univ. Press, 1984) を参照。

- 3) David Crystal, *The Cambridge Encyclopedia of Language* (Cambridge Univ. Press, 1987) を参照。
- 4) Edward Matthei and Thomas Roeper, *Understanding and Producing Speech* (Fontana Paperbacks, 1983) の第2および3章を参照。
- 5) 同上
- 6) E. Lenneberg, *Biological Foundations of Language* (Wiley, New York, 1967) を参照。Adrian Akmajian, Richard A Demers and Robert M. Harnish, eds., *Linguistics: An Introduction to Language and Communication* (The MIT Press) の第6章からの抜粋。
- 7) P. Liberman, "Towards a unified phonetic theory", *Linguistic Inquiry*, 1. を参照。註3の Edward Matthei and Thomas Roeper からの抜粋。
- 8) 註3の Edward Matthei and Thomas Roeper を参照。
- 9) R. Ladefoged and D. Broadbent, "Information conveyed by vowels" (*The Journal of the Acoustic Society of America*, 29, 1957) 註3の Edward Matthei and Thomas Roeper からの孫引き。
- 10) G. Peterson and H. Barney, "Control methods used in a study of the vowels" (*The Journal of the Acoustic Society of America*, 24, 1952)