

チンパンジーの聴覚と音声 ——ヒトの音声言語の理解を目指して——

小嶋 祥三*

要旨：ヒトの音声言語の起源を検討する基礎資料として、系統的に最も近いチンパンジーの聴覚と音声の機能を次の7点から検討した。1. 聴感度と弁別閾, 2. 音声知覚, 3. 母音的音声 grunt の発声, 4. 聴覚認知, 5. 発声訓練, 6. 音声発達, 7. 母子音声交換。

チンパンジーはW字型の聴感度を持ち、周波数、強度分解能はヒトよりも劣っていた。かれらは音声の低い周波数成分を手がかりに知覚していた。かれらは[u], [o] [a]と聞こえる音声 grunt を出した。視覚—聴覚見本合わせの獲得は難しいが可能だった。しかし、名前を理解している証拠は得られなかった。なお、かれらは音声による個体識別の能力は例外的に優れていた。かれらの音声は自発性が乏しく、可塑性が小さかった。訓練で音声の頻度を変化させることは可能だった。訓練の効果は母子間の音声交換にあらわれ、訓練を受けた個体のみが頻繁に自分の子どもと音声交換を行った。

Key Words : 聴覚, 音声, チンパンジー, 音声発達, 母子音声交換

はじめに

チンパンジーの聴覚や音声に対する興味はヒトの音声言語の起源や進化への問いからきている。この問いの最終的な解は得られないので問うても無駄であるという意見もあるが、ヒトが存在する限り発せられる問いである。最終的な解がどうであれ、その時々での見解が表明されてきたし、これからも表明されるだろう。

チンパンジーの言語の研究は、音声言語を用いた Hayes 夫妻の研究が失敗したことをうけて手話や図形語を利用することになり、少なくとも単語レベルの理解や表出はかれらも可能であることを明らかにした。Hayes 夫妻の失敗はチンパンジーの聴覚や音声の研究にブレーキをかけたかもしれない。チンパンジーの聴感度は1930年代に研究があるだけだったし、周波数、強度弁別閾はデータすらなかった。しかし、人工的な手話や図形語の研究の生物学的意味には疑問があるし、何よりもチンパンジーが話さないことを説明しない。そこで、かれらの聴覚と音声の機能を基本的なところから検討することにした。

ろから検討することにした。

1. 聴感度と弁別閾

音を検出するのに必要な反応時間がラウドネスに対応することを利用して、チンパンジーの大きさの等感曲線を得た。図1では各周波数で800msの反応時間を刺激閾としてある。興味深いことに、W字型の曲線となった。同じ方法で測定したヒトはU字型になった。新世界ザル、旧世界ザルはW字型の曲線となるので、チンパンジーはヒトよりもサルに近い聴感度をもっている。ヒトの音声言語が利用している周波数帯域で、感度が低下する。この感度低下の起源を求めてチンパンジーの外耳道の共鳴特性を測定したが、ヒトと同じであった。また、類似のW字型の聴感度をもつニホンザルで蝸牛マイクロフォン電位、聴性脳幹反応 (ABR) などを測定したが、いずれの電位もW字型になった。このことから、感度低

Auditory and vocal functions of the chimpanzee: A search for the origins of human speech.

*慶應義塾大学・文学部 Shozo Kojima : Department of Psychology, Faculty of Letters, Keio University

Sound pressure level (dB)

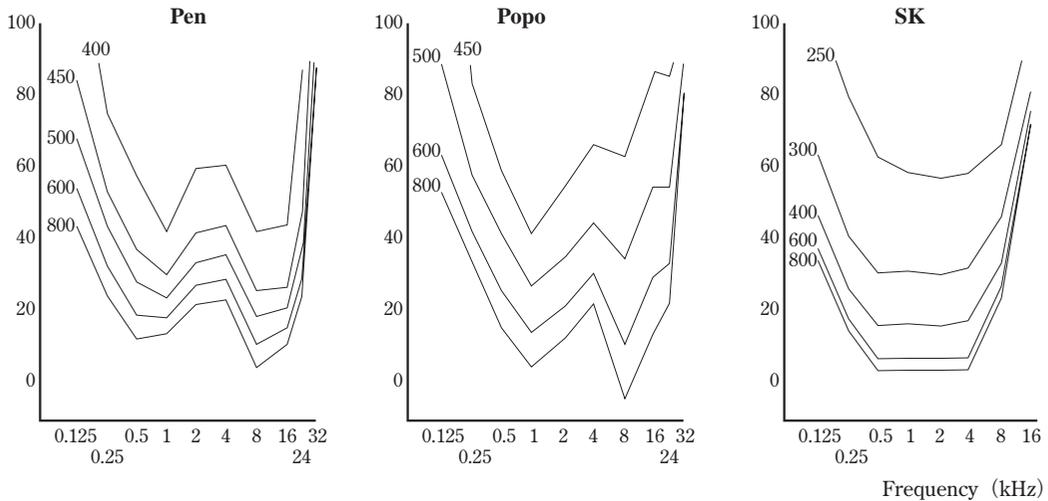


図1 チンパンジー (Pen, Popo) とヒト (SK) のラウドネス曲線。図中の数字は反応時間 (ms)

下の起源は鼓膜から蝸牛の有毛細胞の間にあると推定した。

次に周波数と強度の弁別閾を測定した。1 kHz, 70 dB SPLで比較すると、チンパンジーの周波数弁別閾は約10Hz、強度弁別閾は約1 dB ヒトよりも大きかった。ニホンザルではさらに弁別閾が大である (図2)。音声言語は周波数、強度の微妙な違いで情報を伝えている。ヒトがよい聴覚を持つのは音声言語を持っているからだろう。

2. 音声知覚

音声知覚に関しては母音、子音、チンパンジーの母音の音声 (grunt) の測定を行った。母音では母音間の知覚的類似性、声道長の正規化、プロトタイプ (磁石) 効果を、子音では子音間の知覚的類似性、カテゴリカル知覚、文脈効果を検討した。チンパンジーの grunt の知覚ではかれらにとって重要な周波数成分の同定を行った。ここでは母音の知覚的類似性と grunt の知覚について述べる。2つの母音の弁別に長い反応時間を要するな

らば、その2つの母音は類似して聴こえていると考えられる。図3に日本語5母音のチンパンジーにとっての反応時間に基づく知覚的類似性を多次元尺度構成法で示す。チンパンジーにとって類似して聴こえるのは [u]-[i] と [o]-[e] の対である。これらの対では第1フォルマントが類似しており、第2フォルマントが異なっている。したがって、チンパンジーは第1フォルマントを主要な手がかりにして、母音の弁別を行っていることがわかった。

この点をさらにチンパンジーの grunt ([o] と聴こえる) で検討した。デジタル・フィルターで第1, 2フォルマントを構成している周波数成分 (倍音) を選択的に除去し、もとの声と弁別させた。図4に結果を示す。この図からわかるように、第1フォルマントを構成する基音や第2倍音を除去した刺激は弁別の成績がよい。これらの加工音声刺激はもとの音声と異なって聴こえるようである。逆に第4, 5倍音のように第2フォルマントを構成している成分音を除去すると、言い換えると、基音や第2倍音が残っていると弁別が悪くなった。もとの音声と同じように聴こえるよう

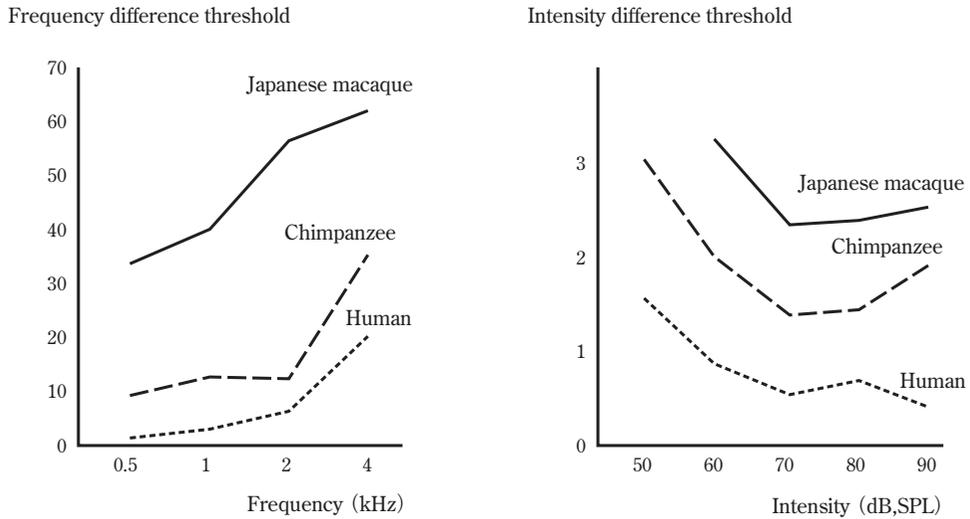


図2 ヒト，チンパンジー，ニホンザルの周波数弁別閾（左）と強度弁別閾（右）。周波数弁別閾は70 dB SPL，強度弁別閾は1 kHzの純音を使用した。

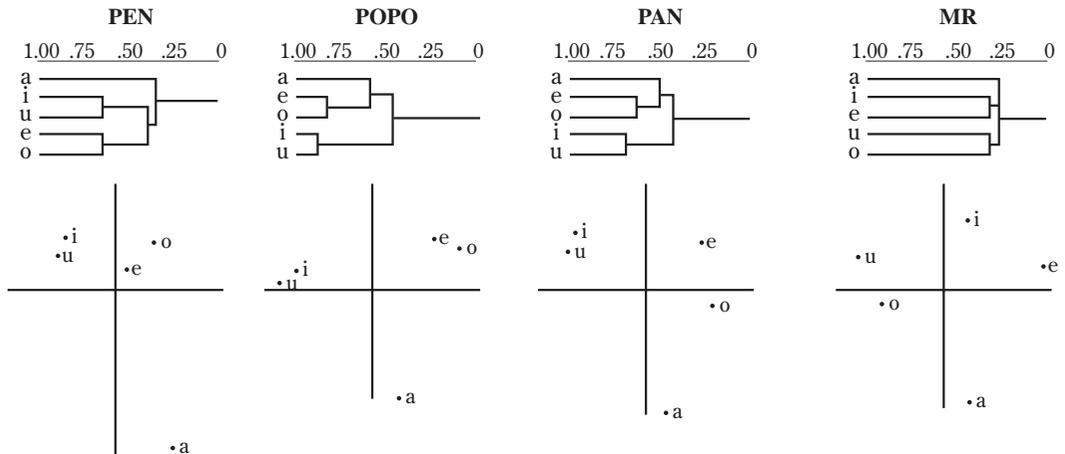


図3 3頭のチンパンジー（PEN, POPO, PAN）とヒト（MR）の合成日本語5母音の知覚的類似性。

である。これらの結果はヒトの母音の知覚の結果と整合的である。チンパンジーは基音など低い周波数成分を聞いている。これは基音が発声者の情動状態など、かれらの社会生活に重要な情報を運んでいるからだろうか。

3. チンパンジーの母音的音声：grunt

以上はチンパンジーの聴くことに関する機能であったが、それと関連するであろう発声の機能について述べる。チンパンジーは様々な音声を発す

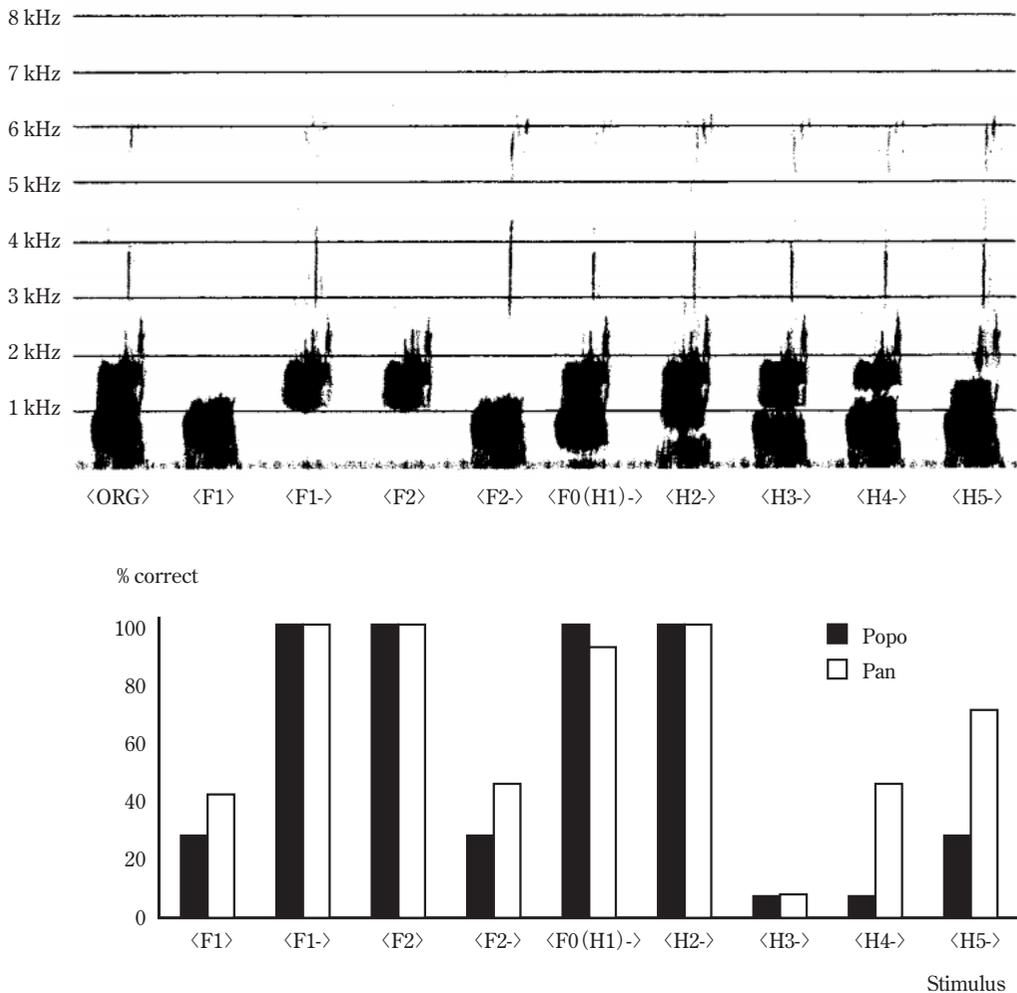


図4 2頭のチンパンジー (Popo, Pan) の母音の音声 grunt の知覚。図で<ORG>はもとの音声, 他はデジタル・フィルターで加工した音声。<F1>第1フォルマントのみを残した音声, <F1->は第1フォルマントのみを除去した音声。第2フォルマント (<F2>, <F2->), 各倍音 (<Hn->) についても同じ。

るが, ここでは不快な情動に関係しない, 非叫喚的で母音的な音声 grunt をとりあげる。1頭のチンパンジーの乳児を人工飼育し, 後述するように発声の訓練をしたが, その間に記録した音声を分析の対象にした。そのチンパンジーは [i], [e] に聞こえる grunt を発することはなかった。[u], [o], [a] に聞こえる音声が多かった。音響分

析をして, 第1, 2フォルマントを測定した結果を図5に示す。第2フォルマントの変化が少なく, 第1フォルマントのみが変化している。日本語の5母音をあわせて表示したが, チンパンジーの grunt は一次元の変化のみで, 母音ラインを形成していた。

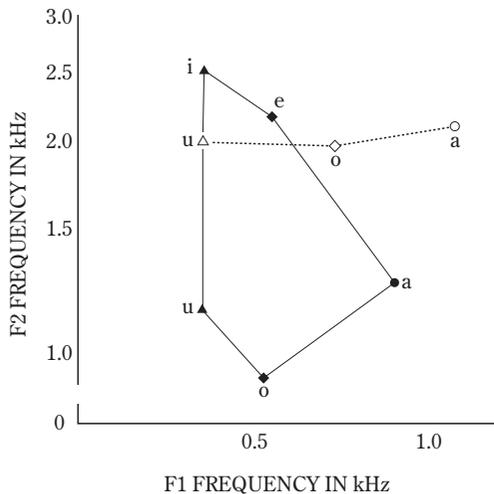


図5 チンパンジー幼児の grunt の F1-F2 図 (点線)。実線は日本語 5 母音。

4. まとめ：聴覚と音声の一致

聴覚と音声が一一致系を形づくることはコロロギからヒトまで多くの種で見られる。チンパンジーの発声の特徴は当然聴取の特性に影響を持つはずだ。すでに述べたように、チンパンジーは第1フォルマントなどの主に低い成分を手がかりに音声を聴取していた。第2フォルマントの変化には鋭敏でなかった。これはかれらの音声は第2フォルマントの変化が少ないことと関係するだろう。また、W字型の聴感度で感度が低下する周波数領域に第2フォルマントがあることと無関係でないだろう。

では、なぜヒト以外の真猿類ではW字型の聴感度になるのか。この問いはまだ答えられていない。1つ考えられるのは、かれらがもつ大きな遠距離交信用の音声である。これは聴取の機能に障害を与える可能性がある。内耳の蝸牛の有毛細胞の保護のためと考えることができるかもしれない。感度低下の起源が有毛細胞よりも末梢、恐らく中耳にあることと整合的だろう。

5. 聴覚認知：視聴覚間マッチング

霊長類の聴覚認知の研究は少ない。その理由は訓練が難しいことにある。任意の音と視覚刺激の連合、マッチングをヒトを除く霊長類は一般に苦手とする。チンパンジーといえども例外でない。2, 3年の試行錯誤を経て、チンパンジーが慣れ親しんだベルなど音だす物体を用いた聴覚—視覚のマッチング (見本あわせ課題) の訓練をした。見本には音を出しながら物体を見せる (AV) 条件、音を出さず物体のみ見せる (V) 条件、物体を見せないで音のみ提示する (A) 条件があり、反応は物体の映像に対してなされた。すなわち、テスト刺激は2つの物体の映像があるが、その一方は提示された見本刺激にマッチし、他方はマッチしない。マッチする方を選択することが正解である。

見慣れた音のでる物体を使用したにもかかわらず、この課題を獲得するにも多くの試行を必要とした。獲得初期における各見本条件での成績と反応時間を図6に示す。この結果からわかるように、AV, V条件では成績がよく、反応時間も短い。A条件では、たとえマッチングの成績がよくても、しばしば非常に反応時間が長くなる。この間、チンパンジーは映像を見ず、天井の方を見上げている。何か思い出そうとしているような印象を持つ。その後、多くの音の出る物体で訓練をした結果、反応時間が極端に遅くなることはなくなった。

チンパンジーは自分の名前を理解しているようなので、その点も検討した。見本刺激は名前前で、たとえば「アイ」と呼ばれる。テスト刺激はチンパンジーの顔写真で、一方はアイ、他方はアキラの顔である。この場合は、アイの写真を選択すれば正解である。しかしこの課題は獲得ができなかった。無論、任意の音と映像のマッチングもできない。では、チンパンジーが出す音声と顔写真のマッチングはどうだろうか。これは図7に示すように、きわめて容易に、第1試行からできた。筆者も音声で個体認識ができる pant hoot という音声でも、また、筆者には個体識別ができない pant grunt, scream という音声でも容易にマッチさせた。2頭の個体が同時に pant hoot を発声し

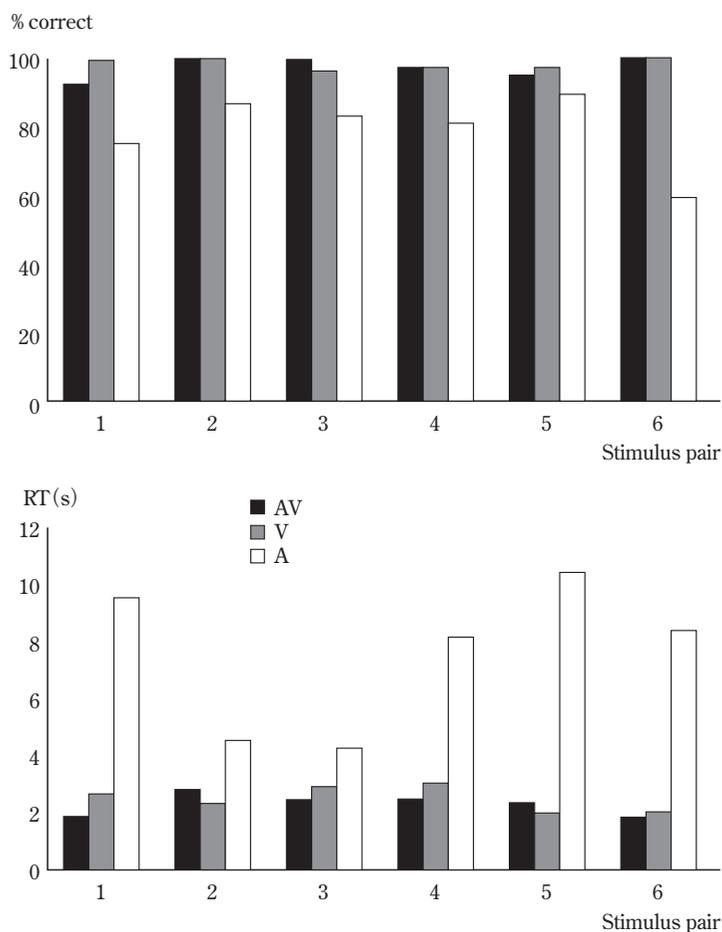


図6 聴覚—視覚のマッチング。上が正反応率，下が反応時間。黒：AV，灰：V，白：A条件。

ても、両方の個体を特定できた。

このような結果をどう解釈すべきか。1つ考えられるのは、訓練が音を出す物体で行われたことである。アイは pant hoot を出すが、「アイ」とはいわない。任意の聴覚—視覚関係のマッチングを訓練されていない。その影響がでたことが考えられる。しかし、音のでる物体、個体のマッチングの訓練、成績、経歴からみて、任意の聴覚—視覚間マッチングへの転移が期待されるが、そのようなことは起こらなかった。ヒト以外の霊長類では異種感覚間のマッチングが困難であるが、その理

由は明らかになっていない。恐らく、彼らにとっては視覚が最も重要なモダリティで、聴覚はそれほど的重要性を持たないのかもしれない。ヒトが音声言語を持ち、素晴らしい聴覚認知機能を持っているのとは大きな違いである。

なお、聴覚—視覚のマッチングで、見本刺激とテスト刺激の間に遅延時間を設けて検討した、チンパンジーの聴覚ワーキングメモリの結果を図8に示す。この図からわかるように、急速な保持の減少がみられた。かれらは音を維持することが難しい。これもかれらの貧弱な聴覚認知の例である。

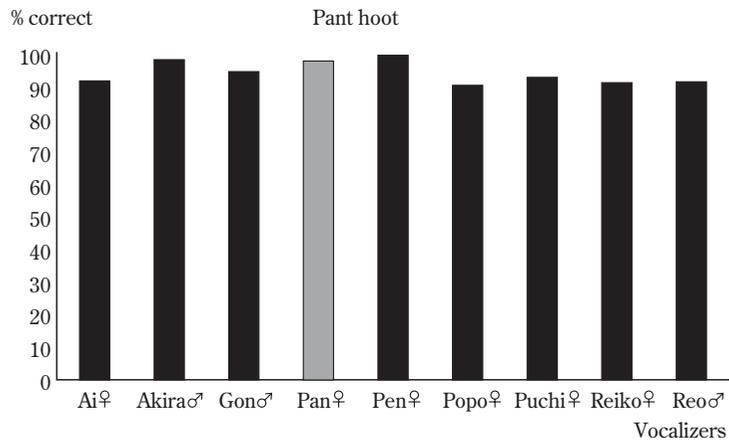


図7 チンパンジーの音声 (pant hoot) による個体同定。灰色は課題を行った被験体 (Pan)。Panは自分の声を理解しているように見えるが、未知の個体を導入すると成績が低下することから、一種の消去法 (exclusion) で反応してことがわかる。それゆえ、音声による自己認知はない。

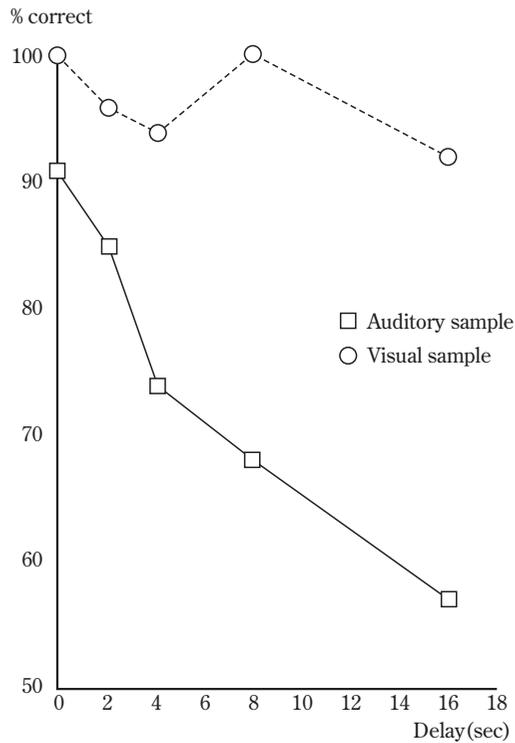


図8 遅延聴覚-視覚マッチングによるチンパンジーのワーキング・メモリの測定。視覚-視覚のマッチングに比べ、保持が急速に減少する。

6. 発声の訓練

チンパンジーに音声言語を教えようとした試みは失敗した。6年間で4語しか「話せ」なかった。声帯振動が伴わなくなり、調音もヒトのそれとは異なっていた。筆者も学習の困難を予測したが、音を単純化して発声訓練を行った。すなわち、バナナには「ア」、ミルクには「オ」と声を出すように訓練した。バナナの結果を図9に示すが、約3週間の訓練でバナナを前にして「ア」と発した。この音声はいろいろな場面で見られるようになり、床に落ちているものに「ア」と発声してから口にすることがあった。この訓練を受けたチンパンジー（パン）は発声頻度が多くなったが、弁別的な発声行動が増えることはなかった。すなわち、いろいろなものに対して異なる音声を発することはなかった。これはヒト以外の霊長類の発声訓練でみられる一般的な現象である。ヒトの音声言語との大きな違いがここにもある。

7. 音声発達

人工飼育の1頭のチンパンジーと母親が育てた1頭のチンパンジーの音声発達を追う機会があった。問題にしたのは非叫喚的音声、staccato, gruntである。両方の個体は共通して、およそ55日齢で発声の頻度を増加させた。その後、しばらく高い頻度を維持していたが、次第に発声の頻度は減少していった（図10）。チンパンジーは興奮していなければ、一般に静かな動物である。特に飼育下で刺激が少ない条件ではこの傾向が強い。

ここで注意しなければいけないのは、チンパンジーの発声が自発的というよりは誘発的な点である。チンパンジーが声を出す確率が高まるのは、ミルクを飲んだ後のユツタリした気分で養育者に注意が向いているときに、低く強い声をかけた場合だった。発声頻度が高い時期は筆者が保育器をのぞくだけで声が出た。その後は見知らぬ人がのぞき込んだときのみ反応するようになった。いずれにせよ、刺激に対する応答として声が出た。これはヒトと大きく異なるところである。Vocal

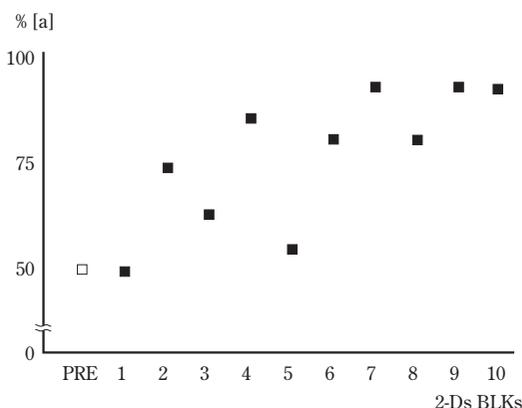


図9 チンパンジーの音声条件づけ。バナナに対して[a]と声を出す訓練。

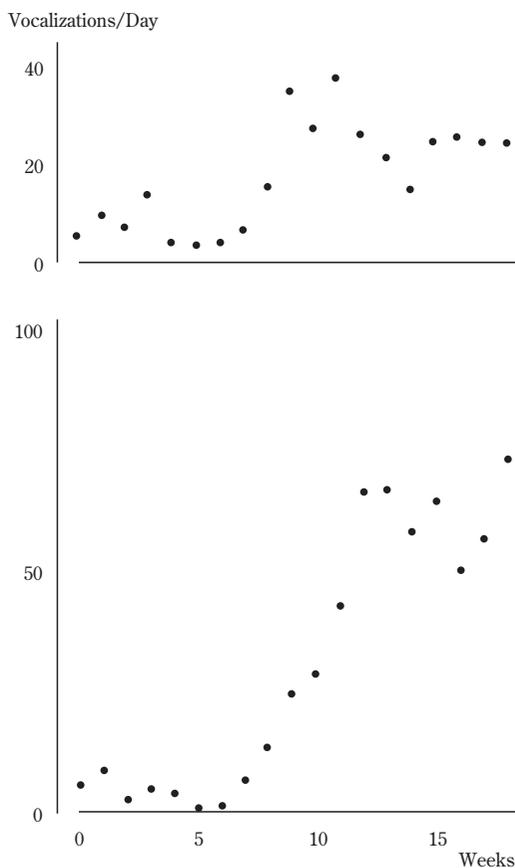


図10 チンパンジー乳幼児の初期音声発達（横軸は週齢）。上はstaccato, 下はgruntの発声数。

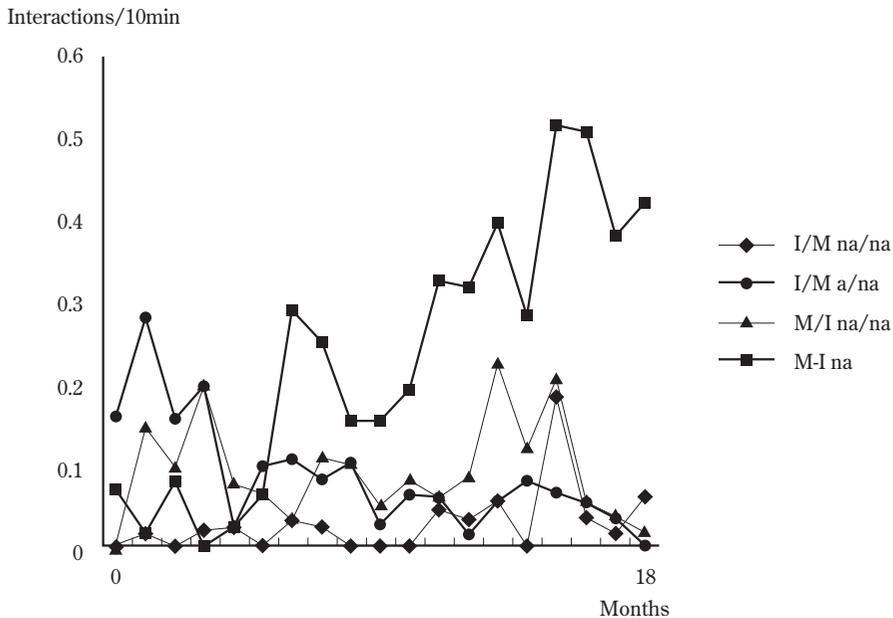


図11 母子音声交換の発達的变化。図でI/M na/naは乳児 (I) が非叫喚的音声 (na) を出し、母親 (M) が同じnaの音声で応えたケース。I/M a/naは乳児が叫喚的音声 (泣き叫び, a) を出し、母親がnaの音声で応えたケース。M-I naは母親が一方的に乳児にnaの音声を出したケース。発達に伴いI/M a/naが減少し、M-I naが増えた。

playというように、ヒトの乳幼児は発声自体を楽しむが、チンパンジーにはそのようなことはない。

もう1つ興味深いことは、チンパンジーは生後0日齢で成体と同じような音声をだすことである。弱いstaccatoは友好的、服従的な音声と思われるが (これは恐らくpant gruntへ移行すると思われる)、0日齢のチンパンジー乳児が発するのを録音した。これらの結果と発声訓練の結果をあわせてかんがえると、チンパンジーの音声は可塑性が少ないように思われる。

8. 母子音声交換

しかしながら、音声の訓練がかねらの生活になんら影響を与えなかったということはない。それ

はかれらと顔をあわせたときの発声量にあらわれたし、以下に述べるように母子間の音声的なインタラクションにも影響がみられた。

母子間の音声交換で興味深いのは、乳児が不快で叫喚的な音声を出し、母親がそれに対してpant gruntのような声で応えるケースである。pant gruntは服従的、友好的な音声なので、母親が「泣く子と地頭には勝てぬ」といった感情を持つかのような印象を与える。この交換は乳児が成長していくにつれ減少したが、母親が一方的に同じ音声で子どもに声をかける行動が増えていった (図11)。この音声交換は発声訓練を受けた個体のみ安定して見られた。

おわりに

以上のように、チンパンジーとヒトは聴覚、音声面で大きく異なる。特に、聴覚認知や音声の可

塑性では相違が大きい。ヒトと異なりかれらは音声言語を持つように生まれついていない。仮にかれらの聴覚や音声の機能が共通祖先のそれを反映しているとするならば、彼らの聴覚や音声の機能がそのままヒトの音声言語に進化したと考えるよりは、例えば身振りを音声言語の前段階と考える方が合理的かもしれない。ヒトの Broca 野は音声の生成と聴取の両者に関わる mirror 性を有し、

それは手の行為にも拡張されることを考えると、言語の身振り起源説は最も無理のない説かもしれない。

文 献

- 1) Kojima, S. : A search for the origins of human speech: Auditory and vocal functions of the chimpanzee. Kyoto University Academic Press, 2003.