

研究題目	脳の高次音声情報処理に関する研究	報告書作成者	宇野宏幸
研究従事者	宇野宏幸		
研究目的	<p>音声認識は実用化段階にあるが、不特定話者認識やノイズに埋もれた音声の認識などには課題が多い。近年、CPUパワーの増大が著しく進み演算スピードが向上しているはいえ、認識アルゴリズムの進化は依然として求められている。これらの要望に答えてゆくためには、洗練された認識アルゴリズムの登場が期待される。</p> <p>本研究の最終的な目的は、生物の脳がもつ音声認識メカニズムを明らかにして、音声認識の工学的開発に結びつけることにある。従来、聴覚処理は末梢神経系に始まり大脳聴覚野で完成すると考えられてきた。ところが、最近の核磁気共鳴装置を用いたヒトの研究は、純粹の聴覚系以外に発声運動中枢が高度の音韻処理に関わっていることを示唆している。なかでも、子音の弁別中に活動性が上昇することから、時間的側面での聴覚処理に不可欠であることが考えられる。これは、音声知覚が発声制御パターンを参照してなされるという「音声知覚の運動理論」を思い起こさせる。このように、発声系が高度に洗練された音声認識マシンである可能性が存在する。しかしながら、ニューロンレベルで神経回路網を明らかにしようとする研究はヒトを用いておこなうことができない。</p> <p>鳴鳥類に分類されるトリは、生後の一定期間に親鳥の音声（歌と呼ばれる）を聞き、聴覚フィードバックを利用して発声を学習する。この過程は、幼児が音声を獲得するプロセスと非常に類似している。なかでも、ジュウシマツ（図1）は歌の各要素をある遷移規則でもって時系列的に配列している（図2）。ジュウシマツは、このような歌要素の系列を認識できるようなものである。このトリの発声系神経核にも聴覚ニューロンが存在している。これらの聴覚ニューロンが、歌の継時的認識に関わっている可能性が十分にある。</p> <p>本研究では、これらの事実をふまえてトリの発声系で実際にどのような神経回路網が形成されて、音声認識がおこなわれているかをしらべ、新しい音声認識のアルゴリズムを提案したい。</p>		

研究内容

ジュウシマツの高次発声制御中枢 (HVC) における聴覚ニューロンの継時的符号化様式とその形成機序を明らかにすることを目的として実験をおこなった。

HVCは、運動制御情報と聴覚情報が統合される神経核である (図3)。HVCは、哺乳類の運動皮質第5層に相当するRAに投射する。RAは、発声器官である鳴管を支配する第12神経核気管鳴管支部 (nXIIts) に線維を送っている。HVCは、また聴覚入力を受ける。大脳聴覚野 (L) の非入力層 (L1、L3) と高次線条体 (HV) からHVCの近傍に線維投射がある。HVCニューロンは、樹状突起を核外に延ばして聴覚入力を受けている。

【方法】被験体は、オスのジュウシマツ (成鳥、孵化後6ヶ月以上)。実験開始前に神経遮断性麻酔薬を投与した。あらかじめ録音しておいた被験体自身の音声 (歌) を提示した時のニューロン活動を同芯円電極を用いて記録した。ニューロン活動は、増幅後にA/D変換されコンピュータ (Macintosh) に取り込まれた。取り込まれたニューロン活動は、複数ニューロン由来のいわゆるマルチプルユニット活動であったが、テンプレートマッチング法を用いてスパイクの分離をおこない、できる限り単一ニューロン活動として取り扱った。

【結果と考察】歌全体の提示中に、ほとんどのニューロンは複数の応答ピークを示した (図4)。HVCニューロンは決まった歌要素に特異的に反応をする傾向は強くない。歌を順方向で再生すると全体的に大きな応答を示したが、時間軸を反対にして再生するとほとんど反応はなかった (図4)。むしろ、抑制される傾向にあった。このように、HVCニューロンは時間方向の情報にきわめて敏感である。ある歌要素を単独で提示した時の応答パターンは、その前に歌要素が存在する場合とでは異なるものであった。歌要素の単独提示では応答が得られない場合でも、この歌要素に時間的に先行する本来の歌要素を配置すると応答が出現するものがあった (図5)。これと比較して、自己の歌の中に出現しない遷移の場合には発火頻度は低い傾向にあった (図6)。応答が出現するためには、1つ前と2つ前の両方の歌要素の存在が必要なものがあった。つまり、3つの歌要素の遷移に関わっているものが見られた。同時に記録したニューロンについて、歌全体に対する応答パターンを比較してみると、少しずつ異なっていた (図7)。これは、近傍のニューロンの継時的符号化様式が微妙に異なっており、これらのニューロンが機能的に結合することによって歌全体の認識に関わっているのかもしれない。

最近、HVCを部分的に破壊したところ、特定の歌要素の遷移が脱落することが判明した (宇野&岡ノ谷 [1998] ジュウシマツ歌の時系列制御、信学技報、sp97-137)。これは、歌要素の発声制御が逐次的におこなわれることを示すが、聴覚符号化においても歌要素間の符号化が同様になされており、両者間にきわめて密接な関係が存在することを示唆する。

<p>研究のポイント</p>	<p>脳の発声系神経核に、高次の音声情報処理機能が期待される。本研究では、この情報処理に関わる神経回路網特性を明らかにする。とくに、発声系での音声処理の特徴と考えられる時間的側面に焦点をあてる。そこで、まずジュウシマツの高次発声制御中枢(HVC)に含まれる聴覚ニューロンがどのように歌要素の系列を符号化しているのかを明らかにする。それらの応答特性がHVC内のニューロン間の相互作用で規定されていることは十分考えられるので、それらニューロン間の機能的結合を明らかにしてゆく。</p>
<p>研究結果</p>	<p>① <u>HVCニューロンは、時間的に側面に敏感であった。</u>自己の歌を正常に再生した場合HVCニューロンは大きな応答を示したのに対して、逆方向再生した時はほとんど応答がなかったかむしろ抑制が見られた。応答パターンは、時間的に先行して別の歌要素が存在する否かで異なったものだった。</p> <p>② <u>歌要素の順番を符号化するニューロンが存在した。</u>歌要素の単独提示では応答が得られないニューロンで、この歌要素に時間的に先行する本来の歌要素を配置すると応答が出現するものがあった。このようなニューロンのなかには、1つ前と2つ前の両方の歌要素の存在が必要なものがあった。つまり、3つの歌要素の遷移に関わっているものが見られた。</p> <p>③ <u>近傍のニューロンの応答パターンは少しずつ異なっていた。</u>ニューロンによって、歌のなかで最大応答を示す要素が異なっていた。</p>
<p>今後の課題</p>	<p>少しずつ応答パターンが異なるニューロンが結合することによって歌全体の認識がおこなわれるのかもしれない。また、ニューロンの時間特性は、細胞間の相互作用によって形成されていると考えられる。そこで細胞間の機能的な入出力関係を吟味することによって神経回路特性を把握したい。このため、トリが獲得した音声をあらかじめ録音しておき、この自己音声を再生した時の神経活動を、多チャンネル微小電極測定システムを用いて脳の発声系から多点で同時に記録する。次いで、相互相関分析法を用いて、神経の発火パターンから神経細胞の相互関係つまり入出力関係を同定したい。</p>

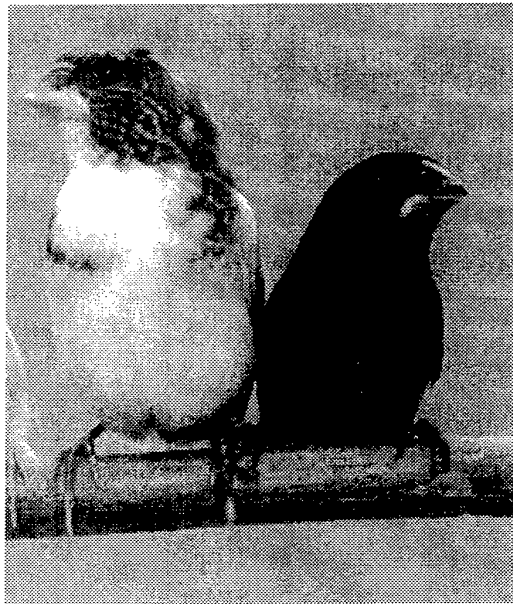


図1 ジュウシマツ。オスのみが音声(歌)を学習する。オスは、歌をうたってメスに求愛する。外観からはオス、メスの区別はつかない。メスは、自分のつがい相手をその歌から識別していると考えられる。

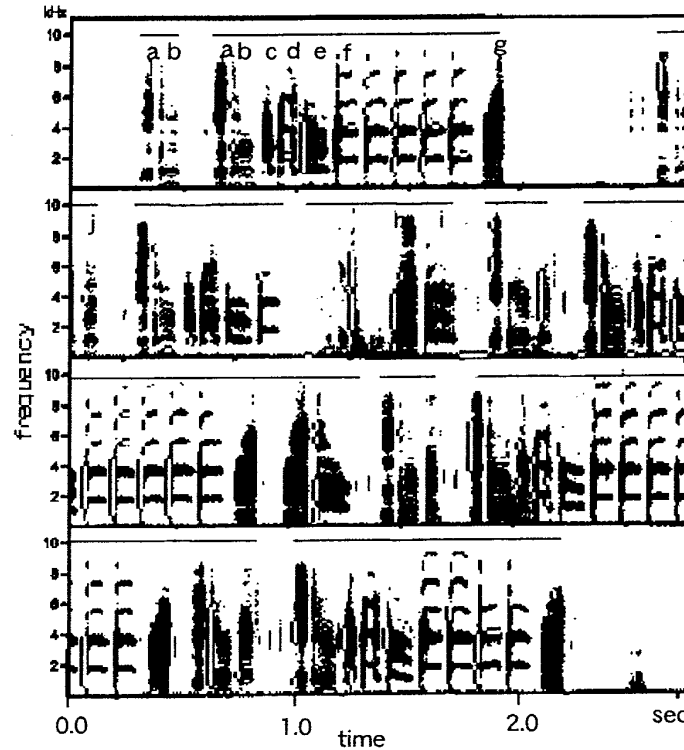


図2 ジュウシマツの歌のサウンドスペクトログラム。各歌要素(a,b,c...)は、無音区間によって隔てられる。ジュウシマツの歌は、これら歌要素がある遷移規則をもって時系列的に配列された構造とみることができる。

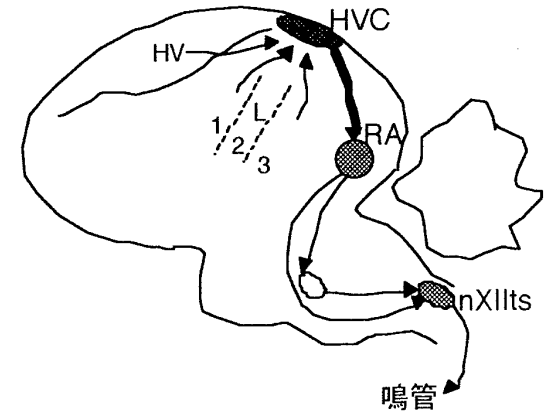


図3 鳴鳥類の発声制御系と高次発声制御中枢(HVC)への聴覚入力。HVCは、哺乳類の運動皮質第5層に相当するRAに投射する。RAは、発声器官の鳴管を支配するnXIIItsに線維を送っている。大脳聴覚野(L)の非入力層(L1、L3)と高次線条体(HV)からHVCの近傍に聴覚入力がある。HVCニューロンは、樹状突起を核外に延ばして聴覚入力を受けている。

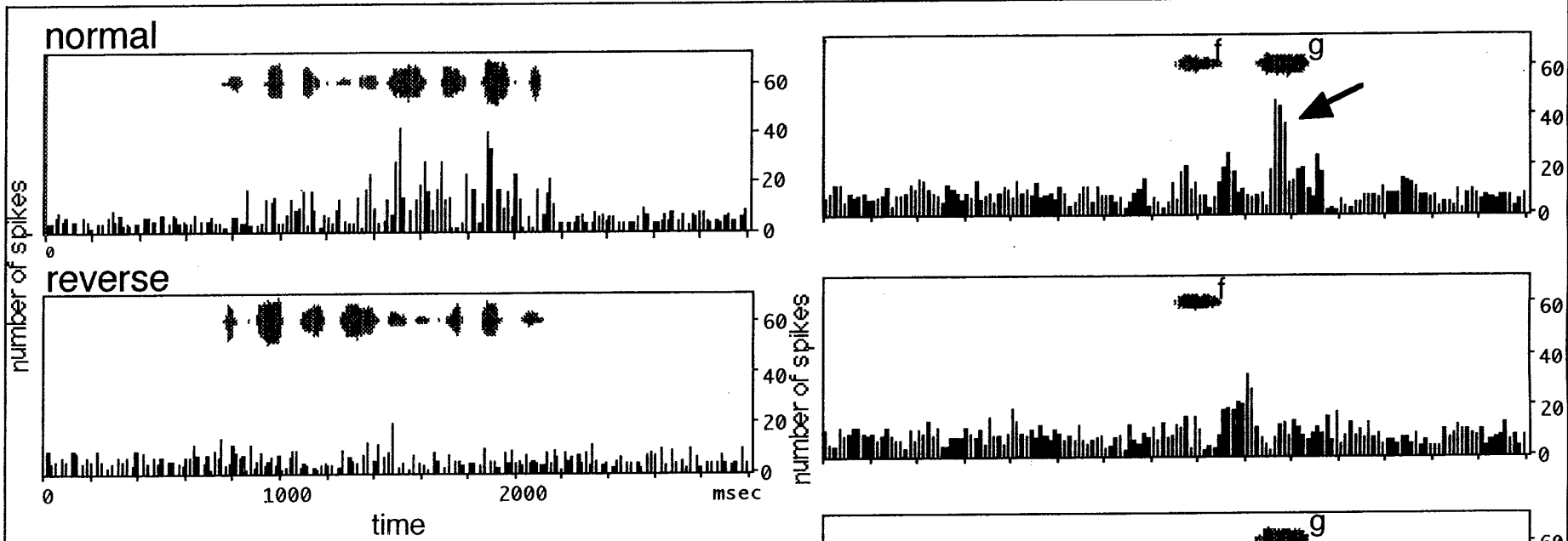


図4 再生方向による応答パターンの相違。通常方向再生 (normal) では大きな複数のピークが認められるのに対して、逆方向再生 (reverse) 時にはほとんど発火はない。各ヒストグラムの上部に刺激波形の振幅包絡を示す。30回提示。ビン幅: 15ms。(以降の図においても提示回数とビン幅は同じ)

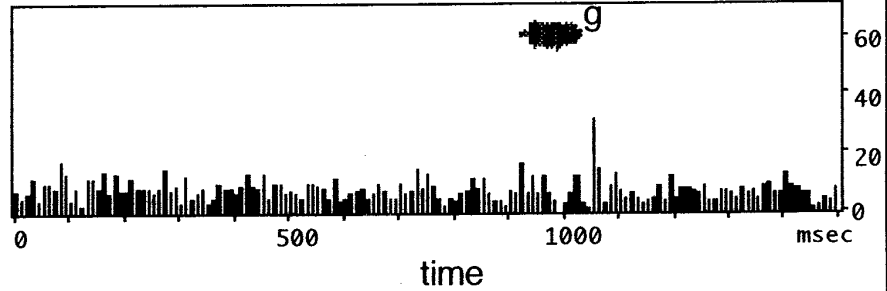


図5 継時的応答の促進。歌中にみられる要素の遷移 f → g を提示すると g の前半部分 (矢印) に対応して発火頻度の上昇が認められる。歌要素 f あるは g を単独で提示しても、当該部位に反応はない。

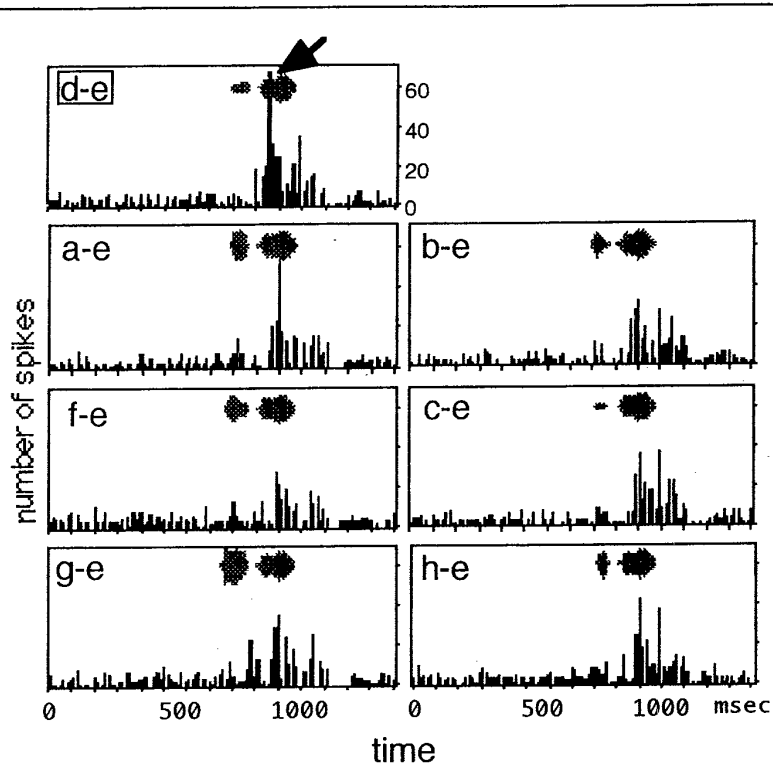


図6 先行する歌要素による応答パターンの相違。歌の中に含まれる遷移d-eを提示すると歌要素eの前半部分で明瞭な一過性応答(矢印)が認められる。これと比較して、歌われない遷移(a-e, f-e, g-e, b-e, c-e, h-e)を提示した時の応答は小さい傾向にある。また、ピーク的位置も異なることに注意。

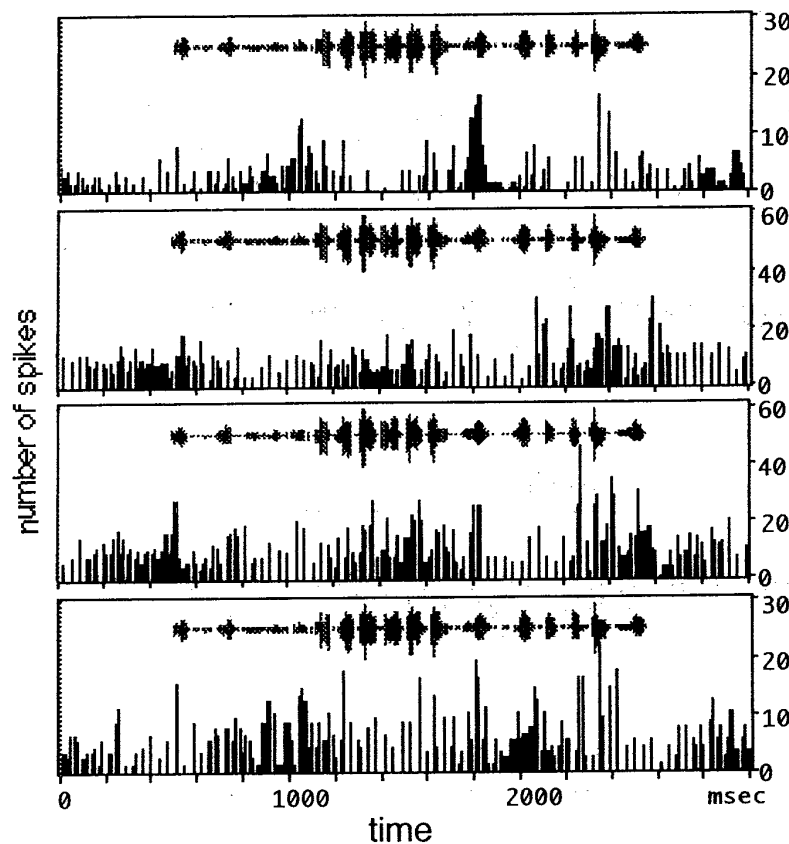


図7 近傍のニューロンにおける応答パターンの相違。同一の電極で記録したスパイクを分類した。いずれも通常方向再生。それぞれの応答パターンは少しずつ異なっている。