

## 第 3 章

### 対連合記憶モデルの検証

#### 3.1 まえがき

第 2 章において, Sakai と Miyashita [3] の生理学的知見を基に, 遅延対連合課題を遂行するモデルを構築した. このモデルは, 生物学的制約条件に反することなく遅延対連合課題の学習及び実行ができるだけでなく, 対連合記憶関連ニューロンの活動をよく再現する. また, 種々の生理学知見を説明すると同時に, TE 野のニューロン活動などに関していくつかの現象を予測する. 但し, これらの予測は, 文献 [3] の実験データでは十分に検証できなかった.

最近, Naya ら [4] は下側頭葉における pair-recall ニューロンの活動をより詳細に測定し, その時間特性を解析することによって非常に興味深いデータを得た. 本章では, 構築した対連合記憶モデルに対して同様の解析を行い, モデルの予測を検証するとともに, 下側頭葉における対連合記憶の形成及び想起のメカニズムについて考察する [31, 32].

#### 3.2 生理学的知見との対応と予測

第 2 章で構成した対連合記憶モデルは, 遅延対連合課題を正しく実行することができ, また下側頭葉の pair-recall ニューロンの活動をよく再現するが, それ以外にもこのモデルには重要な特徴がある. 2.5 で考察した内容のうち, 本章に関係したものを以下にまとめる.

まず,  $N_1$  と  $N_2$  がそれぞれ TE 野と嗅周皮質に相当すると考えると, 図 3.1 のよう

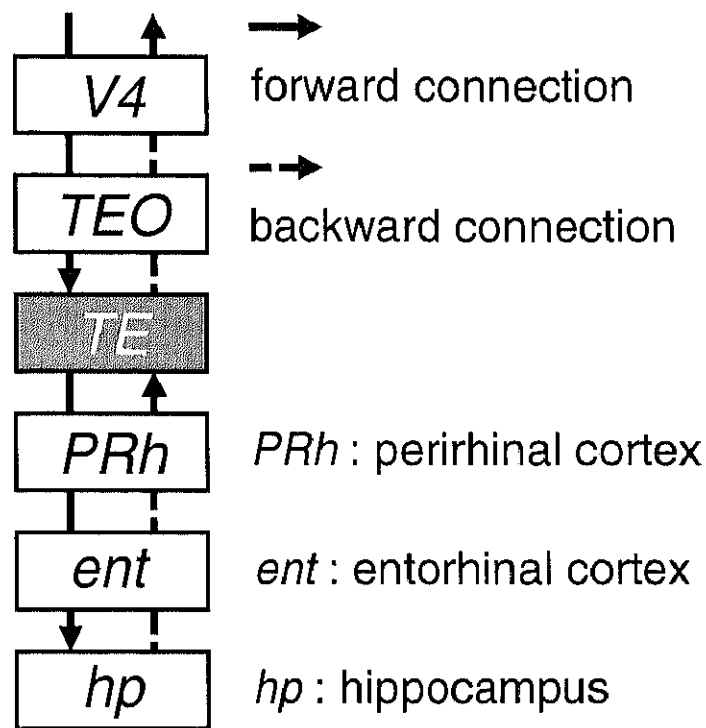


図 3.1: 側頭葉における視覚信号の経路

な解剖学的構造だけでなく、破壊実験の結果とも整合する。例えば、嗅周皮質を破壊したサルは、新たな刺激セットを用いた対連合学習は全くできないが、破壊前に学習した刺激セットに関しては、再学習すれば対連合課題を実行可能である [27]。このことは、嗅周皮質は対連合学習に必要であるが、学習後の対連合課題の実行には不可欠ではないことを示唆しており、このモデルによってうまく説明される。

また、対連合課題における TE 野ニューロンの活動に関して、次のような予測が導かれる。

1. 遅延期間中の活動変化のタイミングは、ニューロン毎に異なっているはずである。このことは、回路網  $N_1$  の状態が跳躍的に変化するのではなく、軌道アトラクタに沿って連続的に遷移することから帰結される。
2. ある図形 A を cue として提示したときのニューロン群の活動パターンと、その対図形 B を cue 提示した後に想起される活動パターンとは少なからず異なる、すなわち遅延期間中に想起されるのは target の完全なコードではないはずである。こ

れは、図 2.9 において a と a' とが離れているからであるが、このことは反対方向に向かう 2 本の軌道アトラクタを相互干渉を抑えて形成するためには避けられない。

以下では、Naya らの実験データを基にこれらの予測を検証する。

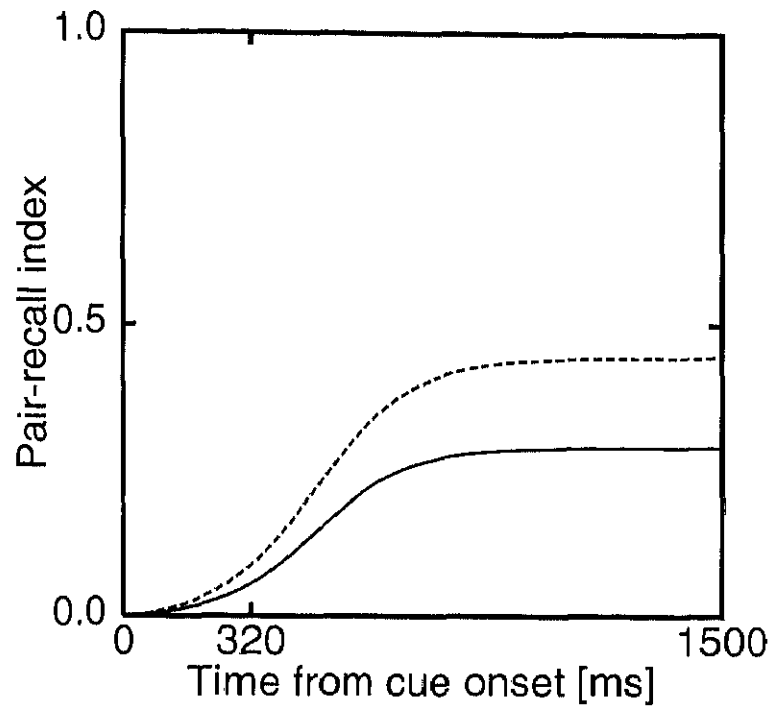
### 3.3 想起の時間特性の解析

まず、Naya ら [4] の解析方法と結果の概略を述べる。彼らは、Sakai と Miyashita [3] と同じ遅延対連合課題についてサルを十分に訓練した後、課題実行時の TE 野及び 36 野 (嗅周皮質の一部で TE 野に隣接する領域) のニューロン活動を計測した。そして遅延期間中に選択的な活動を示したニューロンについて、それぞれ以下の式に従って PRI (pair-recall index) と呼ぶ指標を計算した。

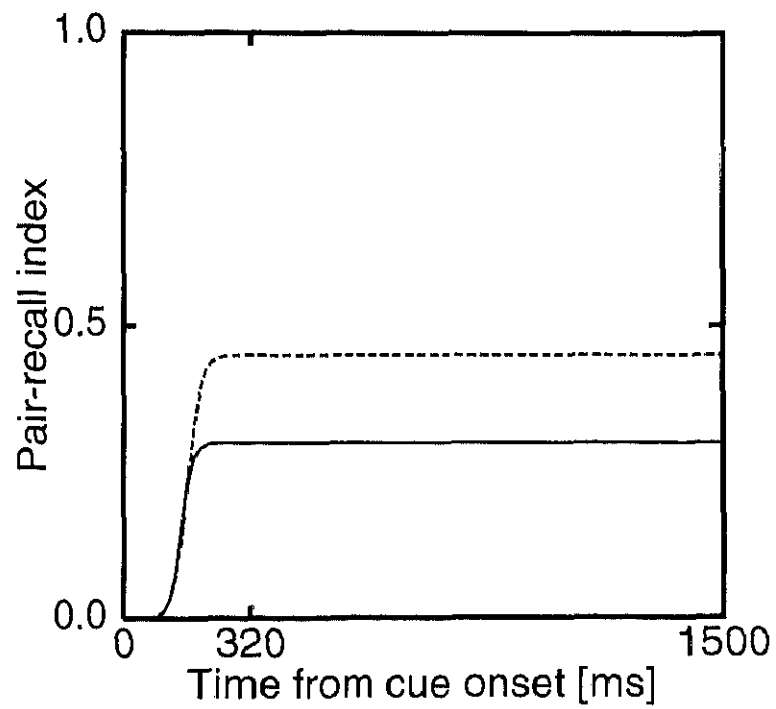
$$PRI(t) = \frac{\langle C_p | F(t) \rangle - \langle C | C_p \rangle \langle C | F(t) \rangle}{\sqrt{(1 - \langle C | F(t) \rangle^2)(1 - \langle C | C_p \rangle^2)}} \quad (3.1)$$

ここで  $F(t) = [f_1(t), \dots, f_l(t)]$ ,  $C = [c_1, \dots, c_l]$ ,  $C_p = [c_{p(1)}, \dots, c_{p(l)}]$  であり、 $l$  は学習させた図形の数 (この実験では  $l = 24$ )、 $f_k(t)$  は  $k$  番目の図形を cue としたときの時刻  $t$  における活動 (発火頻度)、 $c_k$  は  $k$  番目の図形を cue 提示したときの活動、 $c_{p(k)}$  は  $k$  番目の図形の対図形を cue 提示したときの活動、 $\langle X | Y \rangle$  はベクトル  $X$  と  $Y$  の相関係数をそれぞれ表す。この指標は、時刻  $t$  における各ニューロンの反応が、target に対する反応にどの程度近づいているかを表すものであり、最大で 1、cue 期間中は  $PRI(t) \simeq 0$  となるような正規化がなされている。また、cue 期間中と遅延期間中の活動レベルの違いによる影響は受けない。

PRI を計算した TE 野ニューロンについて  $PRI(t)$  を平均し、curve fitting を行った結果が図 3.2(a) のグラフである。実線はすべてのニューロンの平均値、破線は PRI が有意 (5% 水準) に変化したニューロンに関する平均を表す。同様に、図 3.2(b) は 36 野のニューロンに関する平均 PRI を表したものである。ここで注目すべき点は、TE 野における平均 PRI が 500ms 以上にわたって徐々に増加していることである。また、36 野の方が TE 野よりも PRI の増加が早いというのも新奇かつ重要な結果である (図 3.1 からわかるように、cue に対する反応は TE 野の方が早い) が、これについては 3.4 で考察する。



(a) TE 野



(b) 36 野

図 3.2: 下側頭葉 pair-recall ニューロンの平均 PRI の時間変化 (Naya ら [4] より引用)

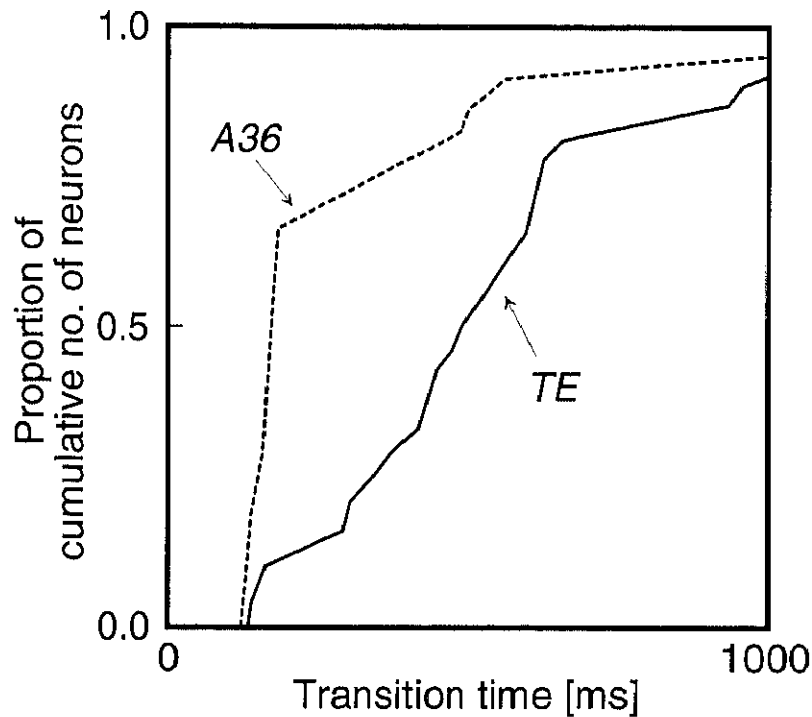


図 3.3: 下側頭葉における TRT の分布 (Naya ら [4] より引用)

次いで彼らは、各ニューロンについて TRT (transition time) と呼ぶ指標を求めた。これは  $PRI(t)$  が最大値の半分まで増加するまでの時間である。図 3.3 は、TRT の累積度数分布を示したもので、実線が TE 野、破線が 36 野を表す。TE 野の累積度数は少しずつゆっくりと増加しているが、このことは TE 野ニューロンの PRI が増加するタイミングがかなり一様に幅広く分布していることを意味する。

さて、以上と同じ解析を、モデルの  $N_1$  のユニットに対して適用した。具体的には、ユニット数  $n = 1000$  のモデルにランダムに作成した 20 組のパターンを入力し学習させた後、対連合課題を実行させ、ランダムに選んだ 100 個のユニットについて  $PRI(t)$  を計算した。シミュレーションの方法やパラメータは、学習時の遅延時間を  $3\tau$  にしたことを除き 2.3.3 および 2.4.4 と同じである。

図 3.2 と同様に平均 PRI の時間変化をプロットした結果が図 3.4 である。但し、横軸の単位は時定数  $\tau$  であり、curve fitting は行っていない。また、TRT の累積度数分布は図 3.5 のようになった。

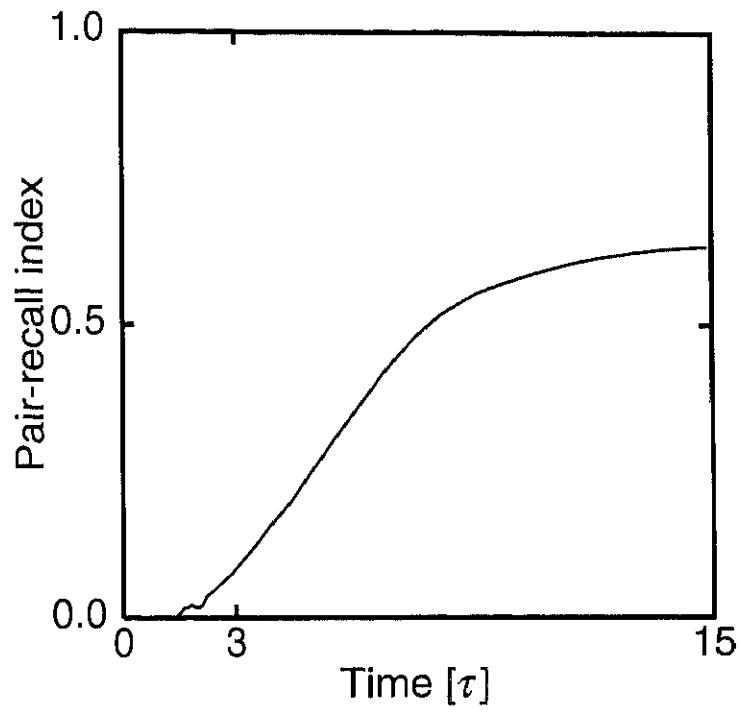


図 3.4: 回路網  $N_1$  のユニットにおける平均 PRI の時間変化

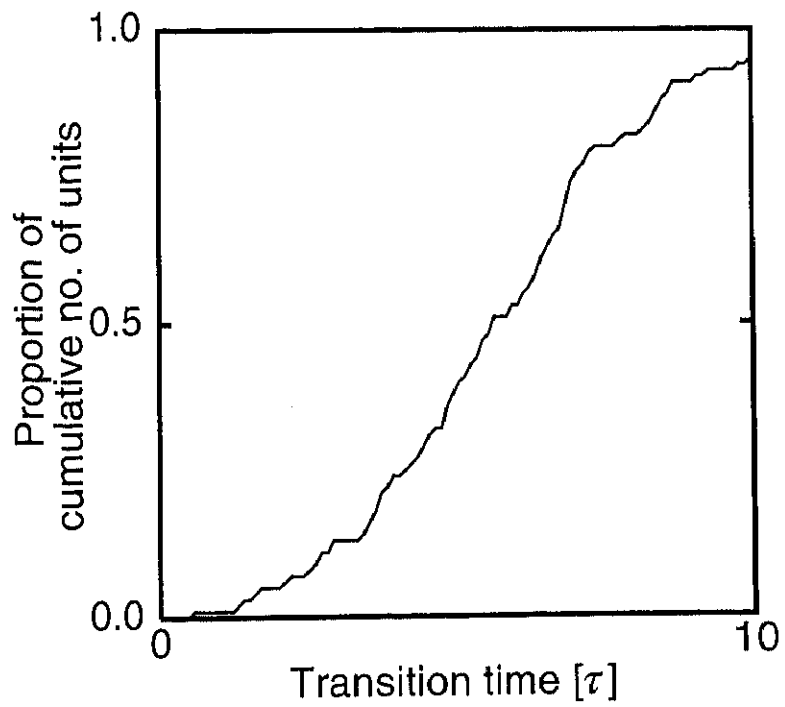


図 3.5: ユニットの TRT の分布

### 3.4 考察

図 3.2(a) と図 3.4 を比較すると、TE 野とモデルの平均 PRI のグラフはよく似た形をしているが、前者の値は後者の半分程度であることがわかる。両者の値が異なる原因として、生理データには target の想起にほとんど寄与しないニューロンが含まれている可能性があるのに対し、モデルではそのようなユニットは全くないことが挙げられる。実際、図 3.2(a) で  $PRI(t)$  の増加が有意ではないニューロンを除いたグラフ (破線) は、定量的にも図 3.4 にかなり近い。

一方、このような操作を行った場合でも、TE 野における遅延期間中の平均 PRI が 0.5 程度までしか増加しないということは、target をコードする活動パターンとその対図形を cue 提示したときに想起されるパターンとがかなり食い違うことを示している。このような食い違いは、test 図形が target に類似していると誤認識を起こしやすいという知見 [30] からも示唆されていたが、Naya らのデータはこれをより直接的に示すものであり、3.2 で述べた予測 2. を実証するものと言えよう。

次に、図 3.3 と図 3.5 を比べると明らかなように、TRT の分布に関してモデルは TE 野のデータと非常によく一致している。実験条件を少し変えてシミュレーションをしても同様な分布となることから、これは単なる偶然ではなく、モデルの動作原理の反映だと考えられる。従って、図 3.3 の生理データは、単に 3.2 の予測 1. と合うというだけでなく、TE 野ニューロン群の活動パターンがモデルと同様な経過をたどって連続的に変化すること、さらにはその変化がモデルと同じメカニズム、すなわち軌道アトラクタによって生じていることを示唆する。

TE 野に軌道アトラクタが形成されているという主張には、もう一つ重要な根拠がある。それは、図 3.2(b) や図 3.3 の破線で示されているように、嗅周皮質 36 野の方が TE 野よりも先に target を予測する活動状態へシフトするということである。前述のように、軌道アトラクタの形成には回路網の状態遷移を先導する学習信号が必要であり、そのような信号が  $N_2$  から  $N_1$  に送られるというのがモデルの骨子の一つである。 $N_2$  が嗅周皮質に対応することを考慮すると、36 野の状態遷移が TE 野に先行するという事実は、嗅周皮質からのフィードバック信号を用いて TE 野に軌道アトラクタが形成されるという仮説を支持する。

ここで、36野の活動によってTE野の状態遷移が引き起こされるのならば、TE野に軌道アトラクタがなくてもよいのではないかという疑問が浮かぶ。確かに、サルでは学習後もTE野の状態遷移が嗅周皮質にもある程度依存している(モデルで言うならば、学習信号の入力強度 $\lambda$ が0でない)可能性が高い。しかし、対連合記憶は嗅周皮質にのみ形成されるという考えは、前述の破壊実験に関する知見と合わない。また、もしTE野には各図形をコードする点アトラクタしかなく、従ってtargetの想起は他の領域からの信号に完全に依存しているならば、一般に別のアトラクタへ遷移する際の過渡状態は不安定であり短時間で終わるから、TRTが図3.3のように広く分布することはないはずである。従って、TE野に軌道アトラクタが形成されているという解釈が最も妥当だと考えられる。

以上のように、TE野のpair-recallニューロンに関する生理データは、対連合記憶の計算論的モデルの予測と非常によく一致した。このことは、TE野に軌道アトラクタが形成されていること、及びそれに必要な学習信号が嗅周皮質から送られていることを強く示唆する。

但し、本研究のモデルは、主にTE野のニューロン活動を説明する目的で構成しているので、回路網 $N_2$ が嗅周皮質の活動を全て説明できるとは言い難い。

実は、学習時の出力を用いて $N_2$ の細胞のPRIやTRTを求めると、図3.2(b)や図3.3の点線と比較的よく似た形のグラフが得られる。しかし、現在のモデルの $N_2$ はtargetを想起するメカニズムを全くもっていないから、PRIが増加するのは $N_2$ にtargetを入力した後である。これに対して、36野ではcue提示の直後からPRIが増え始めているから、 $N_2$ と36野のグラフは横軸の原点が大きく異なる。

従って、36野のニューロン活動を説明するためには、モデルの $N_2$ について生理学的知見を基にした修正が必要である(このことに関しては5.2にて後述)。