

氏名	まき おか しょう ご 牧 岡 省 吾
学位(専攻分野)	博 士 (文 学)
学位記番号	文 博 第 53 号
学位授与の日付	平 成 8 年 3 月 23 日
学位授与の要件	学 位 規 則 第 4 条 第 1 項 該 当
研究科・専攻	文 学 研 究 科 心 理 学 専 攻
学位論文題目	文 字 と 単 語 の 内 部 表 現

論文調査委員	(主 査) 教 授 乾 敏 郎	教 授 清 水 御 代 明	教 授 苧 阪 直 行
--------	--------------------	---------------	-------------

### 論 文 内 容 の 要 旨

本論文では、学習による内部表現の獲得という観点から文字と単語の認知過程を包括的に理解することを目指し、先行研究のデータの再分析とシミュレーション、これまでに提案されたモデルに関する考察、そして心理実験を行っている。

第1章では、文字とランダムドット・パターンの内部表現について検討した。まず、文字同士の混同確率やランダムドット・パターン同士の類似度を測定した先行研究のデータを、多次元尺度法(MDS)を用いて系統的に再分析した。MDSは、刺激同士の心理的類似度に関するデータから、多次元空間における刺激の布置を導出する。アルファベット文字間の混同確率を測定したギルモア(G. C. Gilmore)らの実験のデータをMDSによって分析した結果、幅や丸みなどの大局的な特徴にもとづく布置が得られた。また、ランダムドット・パターン間の類似度を測定したブラウ(D. S. Blough)の実験データから得られた布置も、大きさや規則性といった大局的特徴に対応していた。これらの分析の結果から、文字やランダムドット・パターンのような2次元的な形は、脳内において大局的な特徴の組み合わせによって表現されていると考えられる。このような見方は、神経生理学的な研究から得られた証拠とも一致している。

次に、上記の心理実験で使用された文字やランダムドット・パターンをニューラルネットワークに学習させ、獲得された内部表現を分析した。計算機シミュレーションでは入力パターンの恒等変換をネットワークに学習させた。このようなネットワークが行う変換は主成分分析に近いことが知られている。アルファベット文字の恒等変換を学習したネットワークの内部表現は、心理実験のデータの分析によって導出された布置と類似した構造をもっていた。これは、ネットワークの獲得した内部表現が人間の脳内表現と類似していることを示唆している。ところが、ランダムドット・パターンを入力パターンとして用いた場合には、恒等変換を学習したネットワークの内部表現から得られた布置は、心理実験から得られたものと似ていなかった。この矛盾は入力表現を改良することで解消することができた。以上のようなシミュレーションの結果、入力パターンの恒等変換を学習する単純なニューラルネットワークによって、複数の心理実

験の結果を説明可能であることが分かった。2次元的な形の脳内表現は、主成分分析による情報圧縮という枠組みで統一的に理解することができると考えられた。

第2章では、単語の認知過程に関する生理学・心理学的な研究について概説し、これまでに提案されたニューラルネットワーク・モデルについて考察を加えた。

単語認知過程と関係の深い脳損傷として、純粹失読がある。純粹失読の患者は、文章を読む能力が顕著に損なわれているが、その他の対象を認知する能力や、言語の理解に関する能力は比較的正常に保たれている。このような障害は、複数の対象を同時に認知するためのシステムが損傷を受けることによって生じるのではないかと考えられている。

一方、認知心理学においては、様々な手法を用いて単語認知過程に関する実験が行われてきた。本論文では、単語優位効果と文字の混合エラーという2つの現象に注目して議論を行った。単語や文字を瞬間的に呈示したとき、単語の中に含まれる文字は、単独で呈示された文字よりも認知が促進されることが知られている。これが単語優位効果である。この現象は、単語の表現と文字の表現との間に相互作用が生じていることを示唆している。一方、複数の単語を同時に瞬間呈示した場合には、文字が単語間で混合して見えることがある。これは文字の混合エラーと呼ばれる。どのような条件において文字の混合エラーが生じやすくなるのかを探ることによって、単語の内部表現の性質を探ることができる。以上のような生理学的・心理学的な知見にもとづいて、これまでに提案された単語認知過程に関する3つのニューラルネットワーク・モデルについて解説し、検討を加えた。

第3章では、文字の位置に関する情報の表現に焦点を当てて、モデルの妥当性を検討するための実験を行った。モザー (M. C. Mozer) のモデル BLIRNET においては、文字クラスターユニットと呼ばれる内部表現が用いられている。このユニットは隣接する文字同士的位置関係にもとづいて文字の位置を表現する。実験1では、このような文字位置の表現が実際の単語認知過程で用いられているのかどうかを検討した。被験者の課題は、瞬間的に呈示されるプロープの中にターゲット単語があったかどうかを判断することであった。プロープは2つの単語から構成されていた。ターゲットを構成する2つの文字は、常にプロープの中に含まれていた (e.g. ターゲット: 実用, プロープ: 実証 応用)。これら2つの文字のプロープの中での位置関係を単語内・単語間で独立に操作し、フォールスアラーム率 (実際にはターゲットがなかったのに「あった」と答える反応) に対する効果を検討した。その結果、2つの文字の単語内における位置関係がターゲットとプロープとで同じである場合の方が、異なっている場合より有意にフォールスアラーム率が高かった。これに対して、単語間での位置関係はフォールスアラーム率に影響を与えなかった。これらの結果は、文字の位置が単語内での相対的位置関係にもとづいて表現されていることを示しており、文字クラスターユニットのような内部表現の存在を支持している。

ところで、日本語の文章は、英語などとは異なり、空白によって単語を区切らないで表記される。実験2では、このような表記の特性が単語認知過程に対してどのような影響を与えているのかを検討した。実験2で使用したプロープは、プロープを構成する単語に隣接する空白を、ランダムにひらがなで置き換えたものであった (e.g. あ実証き応用せ)。実験1と同様な手続きで実験を行った結果、実験1と同じく、単語内における位置関係の効果が有意にみられた。これは、漢字とひらがなの区別が文字同士の相対的位

置関係に関する情報を抽出する以前に行われていることを示している。日本語の認知過程においては、漢字とひらがなの境界が、文を単語に分節化するための手がかりとして用いられているのかもしれない。一方、実験1ではみられなかった単語間における位置関係の効果も有意にみられた。これは、複数の単語から構成される刺激全体の大局的特徴を処理する過程が存在することを示唆している。

第4章では、主成分分析による情報圧縮という枠組みが、単語の内部表現の生成過程に対しても適用できるのかどうかについて検討した。さらに、ニューラルネットワークを用いた単語の内部表現に関する研究が、言語という記号系の理解のためにどのような意味を持っているのかについての考察を行った。

## 論文審査の結果の要旨

これまで人間の認知機能に関しては、認知心理学的な手法によりさまざまな特性が調べられてきた。しかしそのほとんどは認知に関する情報処理過程の特性あるいは情報処理容量の限界などに関するものであった。認知機能を明らかにするためには情報処理過程と同時に情報表現に関する研究を進めることが重要である。すなわち外界の対象が脳内のさまざまな処理過程においてどのように表現されているかという問題である。脳内での情報表現を通常脳内表現と呼んでいる。本論文の主題はまさにこの脳内表現に関するものである。脳内表現の研究は、情報処理過程の研究に比べてきわめて難しい。それは脳内表現を実験的に明らかにしようとする場合、これまでの実験事実を検討した上で精密なしっかりとした仮説または思想を持ち、それを確認するための実験を行わなければならないからである。

脳はその目的に対して合理的な情報の表現法を採用しているはずである。対象の脳内表現を作る目的は大きく二つに分けられるであろう。第一は対象物を操作したり、障害物を回避することを可能にするためである。第二は対象の同定および識別である。本論文の研究テーマは文字、単語であるからその目的は識別、同定になる。対象の識別、同定を行うために重要なことは類似性をうまく表現できることである。「醜いアヒルの子の定理」から論理的にはあらゆる対象が同数の可能な述語を持つことが知られている。ある対象Aが対象Bよりも対象Cに類似していると感じるのは脳内において対象の属性に重みを付けているからなのである。つまり対象の認識を行うことは必然的に属性に対する選択、重みづけなどの操作が必要になる。属性に対して取捨選択を行うことは入力情報の次元を縮約して表現していることになる。そこでまず人間がどのように特徴次元を縮約して脳内表現を作り上げているかという事を調べるために、文字やランダムパターンに対して人間が持つ類似度を丹念に調べることから始めた。

類似度を測定する方法には、短時間提示により刺激の同定を困難にし、その結果生ずる混同率から類似度を求めるものや、直接被験者に類似度を評定させるものなどがある。本論文では、主として前者の方法を用いた。混同行列からバイアスと類似度を分離し純粋に類似度だけを取り出す方法がある。論者はこの方法を用いてアルファベット文字やランダムドットなどのいくつかのパターンに対して人間が感ずる類似度を求めた。ついで多次元尺度構成法という統計的手法を用いて、オーディネーションし、類似度空間を求めた。さらに本論文では単に類似度空間を求めたのではなく、その類似度空間がどのような操作によって作り上げられるかを数理的に検討している。

まず人工ニューラルネットワークによって次のような課題を学習させる。人工ニューラルネットワーク

は三層からなり第一層と第三層はそれぞれ入力層および出力層に対応している。第二層のニューロンの数は入出力層すなわち第一層や第三層のニューロンの数よりも少なくなっている。このような構造を持つニューラルネットワークによって心理実験で用いられたアルファベット文字などのパターンを入力し、ニューラルネットワークに入力層の活動をできるだけ忠実に出力層のニューロンから出力できるように学習させる。この学習は収束しその結果、中間層すなわち第二層には圧縮された表現が作りあげられる。

実はこのニューラルネットワークは、統計学で良く知られた主成分分析にほぼ等しい操作を行っていることが知られている。しかし主成分分析との違いは中間層にあるニューロンのポピュレーションによって対象の主成分空間を作り上げているという点である。またそれらが継時的に提示されるパターンを順次学習することによって徐々にその主成分空間を作り上げている点が統計的手法とは異なる。このような学習を行うことによって対象の情報を圧縮できるだけでなく、圧縮された情報から元の信号を再構成(想起)する方法も自動的に形成される。この内部表現は人間がもつ類似度ときわめてよく似ていることがわかった。

第一章の結論は類似度空間は与えられた刺激集合の主成分空間にきわめて類似しているというものである。またそれは簡単なニューラルネットワークの学習によって獲得でき、圧縮された情報から対象のデータを復元することも可能である。類似の考えは以前からあったが、実証的に明らかにしたものは本研究が最初である。

続いて論者は単語の脳内表現に関する研究を行った。二個の単語を同時に提示したとき提示された単語間で構成文字が移動して誤って単語を認知することがある。このような現象は英語で研究されていたが、日本語でもこのような文字の混合現象が生じることを論者は初めて確認した。この混合現象を丹念に調べることによって単語の脳内表現に関する検討を行った。文字の混合に対する位置の効果を実験によって検討した結果、単語を構成する文字間の相対的位置関係が保存された形で混合されることが明らかになった。単語内の文字のコードと位置のコードが独立に存在して単語を表現しているのではなく、それらが一体となった形で表現していると言える。論者は、これまでの文字や単語認知モデルを総合的に検討することによって、一つの具体的な可能性としてウィッケルグレン(W. A. Wickelgren)の三つ組法のような分散表現が有力であることを示唆した。これは、たとえばCATを\_CA, CAT, AT\_の三文字の組(スペースも含む)で表すものである。三つ組み法に限らないが、このような分散表現を脳が採用しているとすれば、混合現象で得られたデータを自然に説明できるのである。このような研究はまたタイプ・トークン問題とも関係しており、今後の研究に大いに期待したい。

一方日本語は、単語を空白によって区切らないことや、複数の系統の文字が混在しているといった特徴を持っている。論者はこれらの特徴に注目して実験したところ、文を単語に分節化するための手がかりとして、漢字とひらがなの境界が用いられていることが示唆された。文字の位置を表現するのにこの境界が使われている可能性も高い。日本語の単語認知に関する従来の研究は、音韻情報に着目するものが多かった。本論文は、形態的なレベルでの処理に注目し、文字や単語に関する脳内での情報表現に焦点を当てた研究としてきわめて高く評価することができる。しかしながら、混合現象の心理実験からは分散表現が示唆されるにすぎず、具体的な日本語特有の表現を明らかにするためには、新たな実験パラダイムを考案す

る必要があると思われる。

以上審査したところにより、本論文は博士（文学）の学位論文として価値あるものと認められる。なお1996年2月15日、調査委員3名が論文内容とそれに関連した事柄について口頭試問を行った結果、合格と認めた。