

■S3 群 (脳・知能・人間) - 2 編 (感覚・知覚・認知の基礎)**13 章 文字・単語の認知記憶**

(執筆者：牧岡省吾) [2008 年 5 月 受領]

■概要■

人間は、ある言語の読みに熟練すると、一度の複数の文字を読み取って単語を認知し、文を読み進めることができる。これを実現するためには、文字の位置、大きさ、変形に関する不変性を実現するとともに、認知した文字を記憶済みの単語とすばやく照合するメカニズムが必要とされる。様々な心理実験により、複数の文字の組み合わせに関する脳内表現が存在し、それらの組み合わせによって単語の視覚的パターンが記憶されていることが示されている。一方、単語の聴覚的認知過程においては、切れ目なく発話される文から単語を切り出し、記憶と照合するメカニズムが必要である。心理実験によって、人間が文を聞くときにはそれまでに入力された音素系列と矛盾しない単語候補がすべて活性化し、入力が進行するにつれて徐々に絞り込まれていくことが分かっている。

【本章の構成】

本章では、まず文字の視覚的認知のメカニズムについて考察し (13-1 節)、次に単語の視覚的認知過程に関する心理学的知見とそれを説明するモデルについて解説する (13-2 節)。更に、単語の聴覚的認知過程に関する心理学的知見とそれを説明するモデルについて解説し (13-3 節)、最後に文字と単語の記憶表現について述べる (13-4 節)。

■S3 群 - 2 編 - 13 章

13-1 文字の視覚的認知

(執筆著：牧岡省吾) [2008年5月 受領]

13-1-1 形の認知と文字認知

文字は平面上に描かれた二次元のパターンである。その意味で、文字の認知は形の認知の特殊なケースである。特定の言語で使用される文字の種類は有限であり、日本語のように多様な文字を使用する言語においても、日常的に使用する文字は数千種類にとどまる。文字の認知は、二次元の視覚パターンを有限のカテゴリに分類する過程である。文字の認知過程には、字体や色などの事例に固有な情報は捨象され、複数の文字の並列的な処理が要求されるという特性がある。

13-1-2 文字の弁別と位置・大きさ・回転に関する不変性

視野内における文字の位置や大きさが変化すると、網膜に映る文字の像は変化し、大脳視覚野への入力パターンも変化する。しかしそのような変化が生じて、人間は文字が変化したとは感じない。これは、脳内の文字認知システムが位置と大きさに関する不変性を実現していることを意味する。このような不変性は、文字だけでなく形の認知一般についても当てはまる。不変性の存在は主観的には自明のことであるが、その実現には複雑なメカニズムが必要となる。一方、文字を回転させた場合、回転角が大きくなると文字の認知が困難になる。これは、回転に関する不変性が限定的にしか実現されていないことを意味する。

13-1-3 文字の視覚的認知のモデル

まず単純なモデルとして鋳型モデルを考えてみよう。このモデルでは、例えば典型的な"A"の形状（鋳型）が脳内に蓄えられており、それを網膜像と照合することによって"A"の認知が成立すると考える。照合を行う前に、上記の不変性が成立している必要がある（**図 13・1**）。具体的には、網膜像における"A"の対応部分に対して位置・大きさ・回転に関する標準化を施した後、脳内における"A"の鋳型との照合を行うことになる。しかしそのような処理を施した後にも、現実の"A"の形は一定ではないという問題が残る。印刷物であっても字体によって形状は大きく異なり、手書き文字であればその変化は更に著しい。

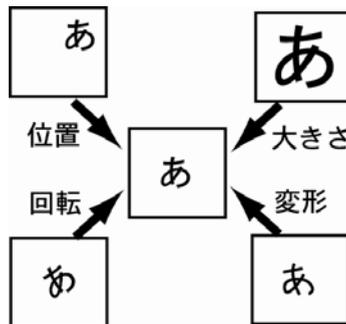


図 13・1 位置・大きさ・回転に関する不変性

そのような変化に対応可能な文字認知のモデルとして、Fukushima のネオコグニトロンがある¹⁾。ネオコグニトロンは、視覚野の神経細胞の振る舞いに関する生理学的知見を踏まえて設計された多層のニューラルネットワークである。大脳一次視覚野には、線分の傾きに対して選択的に反応する細胞が存在することが知られている。そのような特徴検出細胞のなかには、線分が視野内の特定の位置に呈示されたときにだけ反応する単純型細胞と、位置のずれをある程度許容する複雑型細胞がある²⁾。ネオコグニトロンは、単純型細胞をモデル化した S-cell と複雑型細胞をモデル化した C-cell が交互に重ねられた構造をもつ。S-cell は入力パターンに含まれる線分などの特徴を教師なし学習によって抽出し、C-cell は視野内の一定範囲の S-cell の反応を束ねる役割を果たす。文字の変形は視野内で特徴要素をわずかに移動させ、文字位置の変化は特徴要素を大きく移動させる。ネオコグニトロンは、S-cell と C-cell の繰り返しを多段で設けることによって、位置だけでなく変形に関する不変性も実現させる。

一方、Hinton らは、ボトムアップの「認識用」結合だけでなくトップダウンの「生成用」結合をもつようなニューラルネットワークを用いることにより、パターンの変形に対してロバストな認識が可能であることを示した³⁾。彼らは、位置と大きさが標準化された手書き文字パターンを入力として用い、ボトムアップとトップダウンの結合をもつ多層ニューラルネットワークに、文字の分類とイメージ再生(入力パターンの再生)を学習させた。その結果、筆跡のかすれを含むような手書き数字の認識について極めて良好な結果が得られた。これは、トップダウン結合によるイメージ再生により、入力パターンの記述に適した抽象的特徴要素を抽出できたためであると考えられる。人間の文字認知における誤りのパターンが、パターン分類を行うニューラルネットワークより、パターン再生を行うニューラルネットワークのそれに近いことは、Makioka らによっても示されている⁴⁾。

■S3 群 - 2 編 - 13 章

13-2 単語の視覚的認知

(執筆著：牧岡省吾) [2008年5月 受領]

13-2-1 文字認知から単語認知へ

単語は、文字を直線上に並べることによって表記される。単語の意味は、それを構成する文字の種類と順序によって決まる。例えば、「ネコ」と「コネ」は同じ文字から構成されるが、異なる単語である。このように、単語認知では形態情報に加えて順序情報が必要になる。単語認知は、「低次のパターンの順序に基づいて高次のパターンを認知する」という、言語処理に特有の性質もっている。

13-2-2 文脈効果

単語認知過程においては、視野の中心部に呈示された複数の文字が相互に影響を与えることが知られている。これは、ある言語を読むことに習熟した読者は、複数の文字を同時に処理していることを意味している。

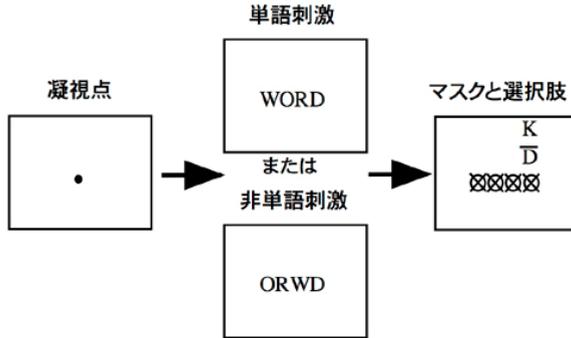


図 13・2 Reicher の単語優位効果に関する実験

文脈効果の一つに単語優位効果がある⁵⁾。これは、単語に含まれる文字の認知が、ランダム文字列に含まれる文字の認知と比べて容易になる現象である。被験者は、瞬間呈示(数十ミリセカンド)された刺激に含まれる文字を、二つの選択肢から選んで答えた(図 13・2)。推測による反応を防ぐため、選択肢中の二つの文字は、残りの3文字と組み合わせた場合に使用頻度のほぼ等しい単語(図中では"WORD"と"WORK")を構成するようになっていた。一方、非単語刺激としては、文字の順序をランダムに入れ替えた文字列が用いられた。実験の結果、単語条件の方が非単語条件より正答率が高いことが分かった。これは、個々の文字の認知が単語に関する知識の影響を受けることを意味する。

13-2-3 単語の視覚的認知のモデル

単語の視覚的認知過程に関する初期のモデルとして代表的なものに、McClelland と Rumelhart の相互作用活性化モデルがある⁶⁾。相互作用活性化モデルはニューラルネットワーク

クであり、特徴レベル・文字レベル・単語レベルの3層からなる。各層に含まれるユニットは、互いに整合性がある場合には興奮性の結合で、整合性がない場合は抑制性の結合で結ばれている。例えば、単語ユニット"TIME"は、文字ユニット"T"からは興奮性の入力を受けるが、文字ユニット"A"からは抑制性の入力を受ける。このような結合を介して各レベルのユニットが相互作用することにより、入力と合致する単語ユニットが最も強く活動し、単語認知が成立する。このモデルは、単語ユニットから文字ユニットからのフィードバック機構に基づいて、単語優位効果を定量的に再現することに成功した。

相互作用活性化モデルの明らかな問題点は、4文字の単語一つしか処理できないことである。Mozer は、ネオコグニロンと同様に位置に関する不変性を徐々に実現していくような構造を用いることにより、単語の文字数や呈示位置の変化に対応することができる単語認知モデル MORSEL を提案した⁷⁾。MORSEL は文字の三つの組に反応するユニットの活動パターンによって単語を表現する。例えば、"WEAR"という単語は、{WE_R, W_AR, EAR, E_R*, AR*, R**}というユニット群の反応によって表現される。このような表現は、単語認知の鍵となる文字の順序情報を保持しつつ、長さの異なる様々な単語を効率よく表現するのに適している。このような順序情報を保持した文字の組み合わせの表現が人間の単語認知過程で実際に使用されていることが、様々な心理実験で示されている^{8,9)}。

■S3 群 - 2 編 - 13 章

13-3 単語の聴覚的認知

(執筆者：牧岡省吾) [2008年5月 受領]

13-3-1 単語の聴覚的認知と語彙獲得

幼児が母語を獲得する際には、聴覚が主要なモダリティとなる(聴覚に障害がある場合は、表記された文字ではなく手話が主要な伝達手段となる)。獲得の初期段階では、まず母語の音声を音素や音節にカテゴリ化することを学習する必要がある。更に、幼児は文を単語へと分節化することを学習しなければならない。これらの学習が語彙獲得の基礎となる。

13-3-2 聴覚的認知過程の時間特性と単語の分節化

Marslen-Wilson は、単語の聴覚的認知過程に関する概念的モデル Cohort を提案した^{10,11)}。Cohort の枠組みは以下の2点に要約できる。①人間が発話を聞くとき、それまでに入力された音素系列と合致する単語の内部表現がすべて候補となり、活性化する。②単語を特定できるだけの音韻・意味情報が入力された時点で、候補が一つに絞られる。

Cohort の枠組みの妥当性は、単語の復唱課題を用いた実験など、多数の心理実験で検証されている。以下に一例をあげる。Marslen-Wilson らは、“captain” (船長) と “captive” (捕虜) のような語頭の音節が共通するプライム刺激を用いて、意味的プライミングの効果について検討した。上記の二つの単語を聴くとき、子音 /n/ の次の母音が呈示された時点で、その単語を一意に特定可能な情報が得られる。実験では、この時点の前後に “ship” (船) や “guard” (見張り) のようなターゲット語が視覚呈示され、被験者はそれが単語かどうかを判断することを求められた(課題を成立させるために一定の割合で非単語も呈示された)。“ship” は “captain” と、“captive” は “guard” と意味的に関連する。ターゲットが子音 /n/ より前に呈示された場合には、“ship” と “guard” の両方において反応時間の短縮が見られた(意味的プライミング効果)。一方、プライム単語が一意に特定可能になってからターゲットを呈示した場合には、意味的に適合する場合 (“captain” → “ship”, “captive” → “guard”) のみ反応時間の短縮が見られた。以上の結果は Cohort モデルの予想を裏づけるものであり、聴覚入力から単語を一意に特定できない時点では、入力と適合する単語候補が同時に活性化しており、特定できるようになった後は候補が絞り込まれていることを意味する。

このように聴覚的認知過程において複数の単語候補が同時に活性化していることは、入力文の分節化と単語認知が不可分であることを意味している。例えば「あかいくつ」という入力を処理するとき、「あか」と「あかい」のどちらを単語候補として認知するのかは、「あか／いくつ」「あかい／くつ」のどちらの分節結果を採用するのかに依存している。つまり、現実的な単語の聴覚的認知のモデルは、分節化過程を考慮したものでなければならない。

13-3-3 単語の聴覚的認知のモデル

Cohort の枠組みに合致するシミュレーションモデルがいくつか提案されている。McClelland と Elman はニューラルネットワークモデル TRACE を提案した¹²⁾。このモデルは、聴覚刺激が呈示されるにつれて複数の単語候補が活性化し、入力と合致する候補が競合に勝ち残る過程を再現することができる。TRACE の特徴の一つは、前述の相互作用活性化モデ

ルと同様、単語レベルから下位レベルへのトップダウン結合をもつことである。音声刺激の一部を曖昧な発音で置き換えたとき、被験者は単語と合致するような発音が聞こえたように感じるという語彙効果 (lexical effect) が知られている¹³⁾。TRACE は、単語レベルからのフィードバックにより、語彙効果を再現することができる。

一方 Norris は、ボトムアップ結合のみで構成されるニューラルネットワークモデル ShortList を提案した¹⁴⁾。ShortList はその名のとおり、単語や単語の一部分に該当する短い音素系列を内部表現としてもつ。単語が呈示されると、入力に一致する内部表現が次々に活性化し、互いに整合性をもつものだけが生き残る。小規模な語彙を対象としたシミュレーションでは、単語候補の競合過程をうまく再現できることが示されている。語彙効果は、ShortList が実在する単語と一致する内部表現しかもたないことによって説明される。ShortList には、TRACE と比べて構造が単純であるという利点がある。

■S3 群 - 2 編 - 13 章

13-4 文字と単語の記憶

(執筆者：牧岡省吾) [2008年5月 受領]

13-4-1 視覚と聴覚のマッピング

我々が単語を音読するとき、視覚情報が音韻情報に変換され、発話される。音読の脳内過程に関しては、現在も論争が続いている。Coltheart らは、音読過程には、単語を媒介せずに文字列を音素に変換する過程と、単語を経て音素に変換する二つの経路が存在するという二重経路モデルを提案した¹⁵⁾。通常の読みでは両方が働くが、初めて学ぶ語や非単語では前者のみが働き、英語の"yacht" (ヨット) など例外的な発音をする語では後者のみが働くとして仮定される。一方、Seidenberg と McClelland は、規則的な発音も例外的な発音も単一のネットワークで扱うことができるトライアングルモデルを提案した¹⁶⁾。このモデルは規則語も例外語も文字から音韻への経路で扱うことができる。この経路は3層のニューラルネットワークで実現され、多数の語の発音を学習することにより、非単語を入力された場合も適切な発音を出力することができる。二重経路モデルとトライアングルモデルのどちらが音読の脳内過程を適切に説明しているのかに関しては現在も論争が続いており、決着をみていない。

13-4-2 文字と単語の脳内表現

視覚系の腹側経路 (ventral stream) において、形に反応する神経細胞の受容野が徐々に大きくなるとともにその選択性が増していく (より複雑な形に反応するようになる) ことが知られている¹⁷⁾。文字と単語の形に関する内部表現が腹側経路に存在することは確実である。Dehaene らは、ニューロイメージングと脳損傷に関する知見と、単語認知のメカニズムに関する考察を総合して、文字と単語の脳内表現に関する枠組みを提案した¹⁸⁾。彼らの枠組みによれば、文字と単語の内部表現は、腹側経路のなかに、徐々に抽象化するかたちで存在する。文字の部分的特徴を検出する細胞は両側の V2 に、形 (字体や大文字/小文字) に依存する文字の検出器は両側の V4 に、抽象的な文字の検出器は両側の V8 に、文字の局所的組み合わせの検出器は左脳の後頭側頭溝 (occipito-temporal sulcus) に、短い語や形態素の検出器は左脳の後頭側頭溝のより前部に存在するとされる。文字の局所的組み合わせの検出器 (local combination detector) は、Mozer が提案した文字の組み合わせを表現するユニット群と同様なものである。この枠組みは、文字と単語の視覚的認知過程に関するこれまでの心理学的・神経科学的研究の成果とうまく整合する。

■参考文献■

- 1) K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol.36, no.4, pp.193-202, 1980.
- 2) D.H. Hubel and T.N. Wiesel, "Functional architecture of macaque monkey visual cortex," *Proceedings of the Royal Society of London, B*, vol.198, pp.1-59, 1977.
- 3) G.E. Hinton and R.R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol.313, pp.504-507, 2006.

- 4) S. Makioka, T. Inui and H. Yamashita, "Internal representation of two-dimensional shape," Perception, vol.25, no.8, pp.949-966, 1996.
- 5) G.M. Reicher, "Perceptual recognition as a function of meaningfulness of stimulus material," Journal of Experimental Psychology, vol.81, pp.274-280, 1969.
- 6) J.L. McClelland and D.E. Rumelhart, "An Interactive activation model of context effect in letter perception: Part 1. On account of basic findings," Psychological Review, vol.88, pp.375-407, 1981.
- 7) M.C. Mozer, "The Perception of Multiple Objects," MIT Press/Bradford books, Cambridge, MA, 1991.
- 8) F. Peressotti and J. Grainger, "The role of letter identity and letter position in orthographic priming," Perception and .Psychophysics, vol.61, pp.691-706, 1999.
- 9) 牧岡省吾, "漢字仮名混じり語における単語優位効果", 神経心理学, vol.16, no.1, pp.66-72, 2000.
- 10) W.D. Marslen-Wilson, "Functional parallelism in spoken word-recognition," Cognition, vol.25, pp.71-102, 1987.
- 11) W.D. Marslen-Wilson and A. Welsh, "Processing interactions and lexical access during word recognition in continuous speech," Cognitive Psychology, vol.10, no.1, pp.29-63, 1978.
- 12) J.L. McClelland and J.L. Elman, "The TRACE model of speech perception," Cognitive Psychology, vol.18, pp.1-86, 1986.
- 13) W.F. Ganong, "Phonetic categorization in auditory perception," Journal of Experimental Psychology: Human Perception and Performance, vol.6, pp.110-125, 1980.
- 14) D.Norris, "Shortlist: A connectionist model of continuous speech recognition," Cognition, vol.52, no.3, pp.189-234 (1994)
- 15) M. Coltheart, B. Curtis, P. Atkins and M. Haller, "Models of reading aloud: Dual-route and parallel-distributed-processing approaches," Psychological Review, vol.100, pp.589-608, 1993.
- 16) M.S. Seidenberg and J.L. McClelland, "A distributed, developmental model of word recognition and naming," Psychological Review, vol.96, no.4, pp.523-568, 1989.
- 17) E.T. Rolls, "Functions of the Primate Temporal Lobe Cortical Visual Areas in Invariant Visual Object and Face Recognition," Neuron, vol.27, no.2, pp.205-218, 2000.
- 18) S. Dehaene, L. Cohen, M. Sigman and F. Vinckier, "The neural code for written words: a proposal," Trends in Cognitive Sciences, vol.9, no.7, pp.335-341, 2005.