

## コネクショニスト・モデルによる漢字非語の音読

伊集院睦雄(いじゅういん むつお)・伏見貴夫・辰巳格  
東京都老人総合研究所 言語・認知部門

(要旨) 漢字語を音読するコネクショニスト・モデルは,頻度・一貫性効果といった単語属性効果をうまく再現するが,未学習の非語を読めないという問題があった.そこで,文字単位の読みに関する知識を獲得しやすい並列入力 系列出力という構造をもつネットワークを構築したところ,ネットワークは単語属性効果を再現し,しかも人間と同程度に非語を音読することができた.この結果は,二重経路モデルと異なる入出力様式に基づく単一のシステムで,単語,非語とも処理が可能であり人間と同様のパフォーマンスを再現できることを示している.

Key words: コネクショニスト・モデル, 漢字非語, 系列処理

日本語の音読研究では,従来,漢字語は単語単位で処理されるといわれるが,文字単位での処理も考えられる.なぜなら,単語頻度効果のみならず,読み方が一通りしかない文字から構成されている一貫語の方が,読み方が複数ある文字から構成されている非一貫語より音読成績が良いという一貫性効果が認められるからである.この効果の説明として,1.文字単位の処理は類似した単語からの類推に基づくとするアナロジー・モデルと,2.単語単位の処理と文字単位の処理に個別のシステムを想定する二重経路モデルがある.我々は,前者のモデルをコネクショニスト・ネットワークで構築し,シミュレーション実験を行ったところ,頻度・一貫性効果を得た.この結果は,漢字語の音読に際し単語単位・文字単位の処理が一つのシステムで実現可能であることを示している.

ところで,文字単位の処理があるならば,非語(例:連料,開員)を読めるはずである.人間は89%の成績で非語を読めるが(Fushimi et al., 1999),我々のネットワークは,7%~20%程度しか読めなかった(Ijuin et al., 1999 など).ネットワークの非語の成績に影響を与える要因として,(i)学習単語数,(ii)入力単語の表現法,(iii)ネットワークの構造,が考えられる.すでに(i)(ii)に関しては検討済みであり,これらを改善するだけでは,人間と同じ成績にまで達しないことが分かっている(伊集院ら,1998,成績は表を参照).

今回,我々は(iii)に注目し,ネットワークの構造を変えることで,非語読みの成績が向上するかを検討した.従来のネットワークでは,単語を構成している文字全てが並列入力されると,単語全体の音韻が並列出力される(parallel-input parallel-output).これでは,非語の読みに必要な,各文字とその読みとの対応に関する知識を獲得しにくい.しかし,単語の全ての文字が並列入力されても,一時刻目には一文字目の読み,次の時刻には二文字目の読みというように,音韻情報を系列出力させることができれば(parallel-input serial-output),

文字単位の読みに関する知識を獲得しやすく,非語読みの成績が向上すると予測できる.

### 1. ネットワークの構築

学習単語は,岩波国語辞典第四版(1993)の漢字2字の見出し語,約31,000語中,出現頻度(国立国語研究所,1970)が5以上で,同字異音語(工場:/koujou/,/kouba/)がない4,120語である.

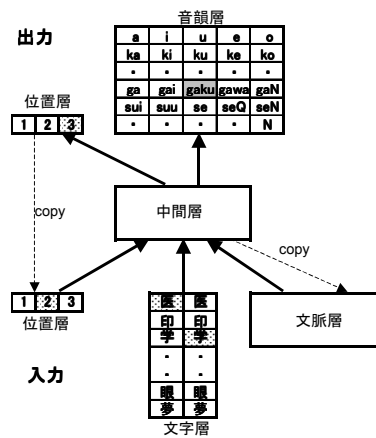


図1 SRNの構造

「医学」を入力した際の第二時刻目の出力状況.入力側の位置,音韻層の活性状態は,第一時刻目の出力側の各層における活性状態のコピー,文脈層の活性状態は,第一時刻目の中間層における活性状態のコピーである.

並列入力 系列出力を実現するため,Elman(1990)のsimple recurrent network(SRN)を採用した.図1に,今回用いたSRNの構造を示す.このネットワークは,位置別に局所表現された文字列が入力されると,その音韻を系列的に出力する.文字層には,各文字位置を表すスロット毎に漢字1,490文字(学習単語4,120語の各文字位置に出現する漢字の総数)

に相当するユニットがあり、これが 2 スロット、計 2,980 ユニットから成る。これらすべての入力ユニットは、100 個の中間層ユニットと結合している。また全ての中間ユニットは、各文字の読みを表す 495 個の音韻層ユニットと結合している。

ネットワークには、これらのユニット群の他に、適切な音韻を順序どおりに出力できるように、文字列中の処理位置を表すための 3 つの位置ユニットが、入力側と出力側の両方に用意されている。入力側の位置層では、ネットワークが現在処理する位置を第 1, 第 2 ユニットで表現し、出力側の位置層では、次に処理しなければならない文字位置を第 2, 第 3 ユニットで表す（第 3 ユニットは、読みの終了を意味する）。これらのユニットも全て中間層ユニットと結合している。入力側の位置層と文脈層には、それぞれ一時刻前の出力側の位置層、中間層における活性状態のコピーが渡される。この構造により、本ネットワークは、並列的な入力に対して、系列的な出力をすることができる。

## 2. 結果と考察

600 回学習した SRN は、学習単語に対し、99.9% の正答率を示し、人間の音読実験 (Fushimi et al., 1999) で用いられた 120 単語に関して、SRN の出力誤差に頻度・一貫性効果がみられた。また、本 SRN に 120 非語を提示した結果、表に示すよう正答率は 85.0% となり、ネットワークの構造を変化させることによって、予想通り音読成績が向上した。

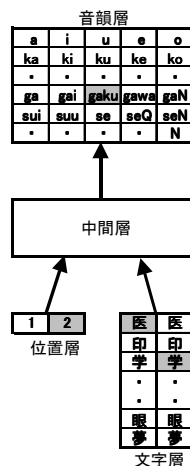


図 2 フィードフォワード・ネットワークの構造  
「医学」を入力した際の二文字目の出力状況。位置ユニットに従って、どちらの文字の読みを出力するかが決定される。

しかし、並列入力・系列出力が問題解決の本質ならば、ネットワーク構造が必ずしも SRN である必要はない。図 2 のような、文字位置

を指定し、系列的に読みを出力するフィードフォワード・ネットワーク (feedforward network: 以下 FFN) でも、良好な音読成績が期待できる。先と同じ 4,120 語を学習させた FFN は、人間の音読実験で用いられた 120 単語に関して、出力誤差に頻度・一貫性効果がみられた。さらに 120 非語を提示した際の正答率は、表に示すよう 91.7% となり、人間や SRN と同程度の成績を得た。

これら二つのネットワークは、並列入力により、一文字目の読みを決定する際に、二文字目の情報を利用できる。よって、一文字目の読み方が複数ある非一貫語を入力した場合でも、「歌手」の出力は/ka/、「歌声」の出力は/uta/と正しく読むことができる。しかも、ある単語 (例: 歌手) の学習は、学習単語中に、その単語と同じ文字を共有し、かつ同じ発音をする単語 (例: 歌謡, 歌曲, 拍手, 投手) が多いほど促進され、一貫性効果が生じる。また、系列出力により、一度の計算では、単語内の一方の文字に対応する読みしか出力しないため、文字単位の読みに関する知識が獲得しやすく、非語の読みを可能にする。

今回の SRN と FFN の処理形態は、明らかに二重経路モデルと異なる。二重経路モデルでは、並列入力・並列出力 (語彙経路) と系列入力・系列出力 (非語彙経路) という二つの処理システムの相互作用によって読みが決定されるからである。本結果は、並列入力・系列出力という入出力様式に基づく単一システムのネットワークで、単語、非語とも処理が可能であり、人間と同様のパフォーマンスを再現できることを示している。位置層と文脈層から中間層への情報が、SRN や FFN の処理にどう寄与しているかに関する詳細は、現在解析中である。

表 各ネットワークにおける非語音読成績の比較

	正答率(%)	構造	入出力様式	学習単語数	表現
Ijuin et al.(1999)	7.5	FFN	並列入力→並列出力	4,136	分散
伊集院ら (1998 Sim.2)	47.5	FFN	並列入力→並列出力	27,751	分散
伊集院ら (1998 Sim.3)	65.0	FFN	並列入力→並列出力	4,136	局所
本SRN	85.0	SRN	並列入力→系列出力	4,120	局所
本FFN	91.7	FFN	並列入力→系列出力	4,120	局所
人間(Fushimi et al., 1999)	88.8				

注: Ijuin et al. (1999) と、伊集院ら (1998: Simulation3) の学習単語には同字異音語が含まれているため、本ネットワークと単語数が異なる。

## 参考文献

Elman (1990). *Cognitive Science*, 14, 179-211.  
 Fushimi et al. (1999). *JEP:HPP*, 25, 382-407.  
 Ijuin et al. (1999). *Psychologia*, 42, 267-280.  
 伊集院ら (1998). 日本心理学会第 62 回発表論文集, 668.