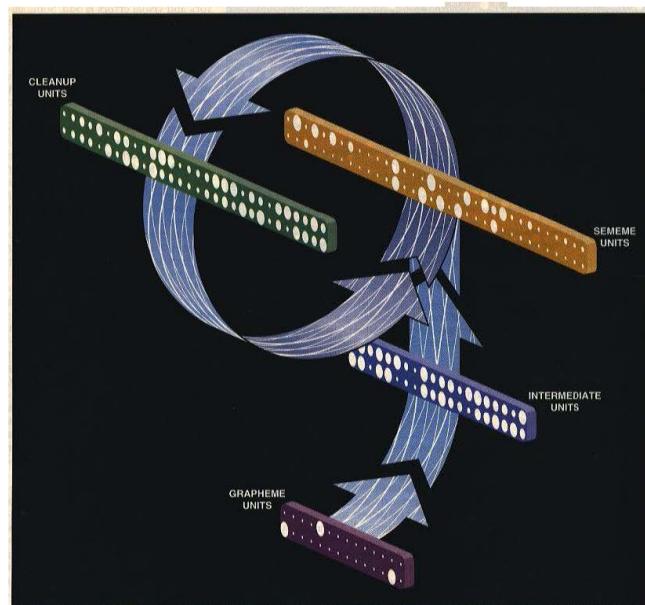


year: 2018 author: Mutsuo Ijuin and Takeko Wydell title: "A Reading Model from the Perspective of Japanese Orthography: Connectionist Approach to the Hypothesis of Granularity and Transparency"

浅川の無駄な補足 Ijuin and Wydell (2018), Harm and Seidenberg (1999) で用いられているアトラクタネットワークの現代的意味について

アトラクタネットワークの起源は、論文にも書かれているとおり Hinton and Shallice (1991), Plaut and Shallice (1993) です。一般には、解説論文 [HintonPlautShallice_SciAme] が良い解説です。



読字のニューラルネットワークには 4 つの層がある。最初の層は、各単語の文字に反応する。入力ユニットと中間ユニット、中間ユニットと「意味素」ユニットの間の結合により、単語の形が、大きさ、食べやすさ、新鮮さなどの意味的特徴の観点からの表現に変換される。「掃除」ユニットは意味素ユニットに結合し、フィードバックループにより、意味素の出力が単語の意味と正確に一致するように調整される。

意味表現を掃除層で圧縮して表現することになるので、後に、**自己符号化器 auto-encoders** と呼ばれるようになりました [Hinton_reducing_dimensionality]。[Hinton_reducing_dimensionality] では、制限付きボルツマンマシンを多層に重ねて、かつ、最後に **ファインチューニング fine tuning** を行うことにより多層化の先鞭をつけたことになります。現在では、このような方法よりも、**一気通貫学習 end-to-end learning** が主流です。

意味表現を、掃除層で圧縮表現することは、**砂時計モデル sandglass models** と呼ばれ、情報圧縮に用いられてきました。ですので、**主成分分析 PCA** との関連や **tSNE** [tSNE] との関連も指摘することができます。

日本語正書法の観点から見た読解モデル：コネクショニストアプローチによる粒状性と透明性の仮説の提唱

• Ijuin & Wydell (2018)

日本人の発達性ディスレクシアの有病率

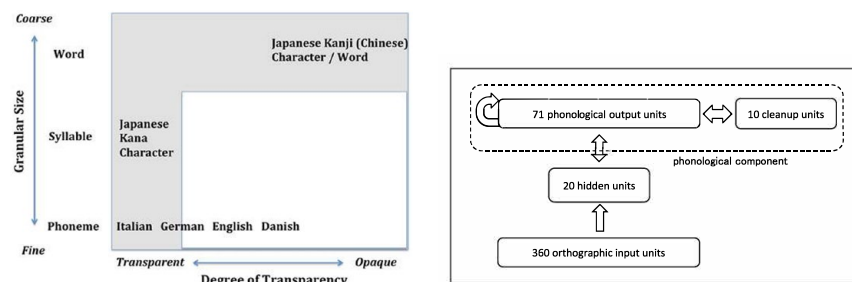
日本人の読字障害（ディスレクシア）の発生率は 0.1%（牧田 1968）や 2 %以下(国立特殊教育総合研究所1996) など、低いことがしばしば報告されている。しかし、これらの研究は、一般的にアンケート調査を用いて行われており、客観的な指標に欠けた。そこで Uno, Wydell, Haruhara, Kaneko, and Shinya (2009) は、日本の小学生 495 人を対象に、読み書きや音韻認識を含むその他の認知能力について調査を行った。その結果、ひらがな、カタカナ、形態素漢字の読解困難者の割合は、それぞれ 0.2 %, 1.4 %, 6.9 %であり、英語の研究で報告されている割合（10%~12%）と比較して、有意に低いことがわかった。日本の研究者は、これらの読解障害を音韻処理の問題ではなく、視覚的または視空間的な問題に起因すると考えている（例えば Kaneko et al.1997）。中国語（もう一つの形態素正書法）で同様の知見を得た Wei ら（2014）は、正書法、音韻法、形態素正書法の認識スキルが中国の小学生の読解成功を予測するものの、正書法の認識スキルが中国語の読解において最も支配的な役割を果たすことを明らかにした。

方法

ネットワークの構成

このネットワークの役割は、かなと漢字の読みを、文字から直接計算することである。図 3 は Harm and Seidenberg (1999) が開発したアーキテクチャをベースにしたネットワークのアーキテクチャである。

今回の実装では、ネットワークの入力層は、かなや漢字の文字ごとに 1 つずつ、360 個の正書法ユニットの集合とした。この正書法ユニットは、ひらがなユニット 71 個、カタカナユニット 71 個、漢字ユニット 218 個で構成された。これらは 20 個の隠れユニットからなる中間層に全結合されており、さらにその中間層は音韻層に全結合されていた。音韻層では、各文字の発音をモーラの並びで表現した。音韻層は 71 個の音韻ユニットで構成され、各モーラに 1 つずつ対応した。各音韻ユニットは、自分自身を含まない他のすべての音韻ユニットに接続された。さらに、これらの音韻ユニットは 10 個のクリーンアップユニットにも接続されており、これらのクリーンアップユニットは、音韻ユニットからの接続を受けたり、音韻ユニットに接続を送ったりする。これらのクリーンアップユニットは、音韻ユニットから活性化を受け、音韻ユニットに活性化を送り返すという、音韻ユニットとの双方向の関係を持ち、クリーンアップユニットによって、音韻ユニット間の直接的な接続だけでは得られない、より高次の音韻依存性の符号化が可能になる (Hinton & Shallice, 1991; Plaut & Shallice, 1993)。今回のシミュレーションでは Harm and Seidenberg (1999) に倣い、音韻層、クリーンアップ層、両層間の重みをまとめて「音韻成分」とした。



左: Fig.2, 右: Fig.3

仮名と漢字の違いは、今回のシミュレーションでは、仮名は 1 文字あたり 1.0 モーラ、漢字は 2.0 モーラである。HGT (Hypothesis of Granularity and Transparency) の観点から見ると、仮名と漢字の透明度は同じ（どちらも発音は一つ）だが、粒度が異なる。

訓練コーパス

訓練コーパスとして、かな 142 文字（ひらがな 71 文字、カタカナ 71 文字）、漢字 218 文字を使用した。このコーパスでは 1 つの仮名文字が日本語の話言葉の 1 モーラに対応し 1 つの漢字文字が 1~4 モーラ（1 モーラ 37 文字、2 モーラ 145 文字、3 モーラ 35 文字、4 モーラ 1 文字）で構成された。すべての仮名文字は、現代仮名遣いの文字で構成されており、短縮音（例 /kyo/）や逸脱子音（例 /kitte/）に対応する文字を除いている。漢字 218 字は、以下のような特徴を持つ：

1. すべて訓読みの言葉であるため、各文字の発音は 1 つ。
2. 漢字の習得年齢 (Amano & Kondo, 1999) が小学校 6 年生 (12 歳) 以下（前述のように、日本の小学生は学年ごとに異なる漢字セットが導入されている）。
3. 仮名文字と同様に、短縮音や geminate 子音に対応する漢字は除外する。
4. 同じモーラの繰り返しのある漢字は除外する（例：母/ha-ha/, 心/ko-ko-ro/）。
5. 構成するモーラの発音が、モーラの順序を入れ替えることで別の単語になる漢字は除く（例：坂/sa-ka/, → 傘/ka-sa/）。

訓練方法

このネットワークは Harm and Seidenberg (1999) が開発したものと同一訓練方法を使用した。実際の人間の子供のように、ネットワークは、正書法の情報が導入される前に、まず音韻の知識を獲得した。つまり、図 3 の音韻構成要素は、外部入力がない状態で、ターゲットとなる発音の音韻パターンを保持するように訓練された。この音韻課題では、訓練コーパスに含まれる 360 文字すべての発音を用いて音韻コンポーネントを訓練した。

健常読字モデル

音韻構成要素が音韻知識で訓練された後、一連の隠れユニットを介して正字入力ユニットを音韻構成要素に接続することが行われ、その結果、すべての漢字と仮名文字を読むために読解ネットワーク全体（図3）が訓練された。Harm and Seidenberg (1999) に倣い、読解学習課題にインターリーブ学習法を導入した。読解課題間に、ネットワークは音韻課題の訓練も受けた。どの課題でネットワークを学習するかは、乱数によって決定された。今回のシミュレーションでは、80 %の試行で読解課題を、20 %の試行で音韻保持課題を学習した。

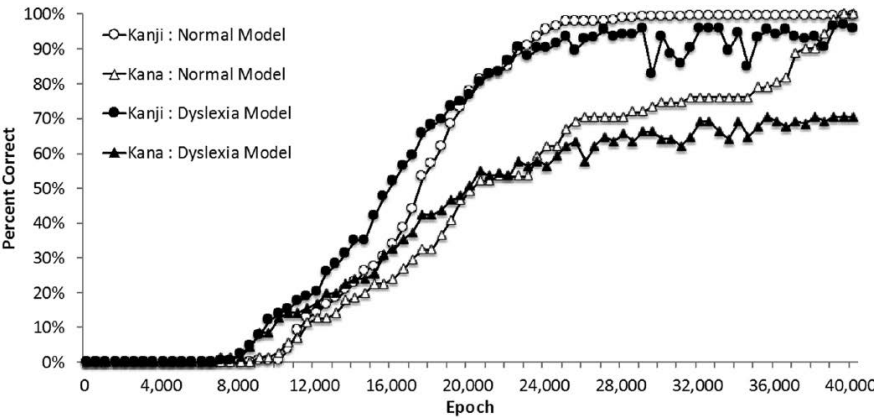
本シミュレーションでは、文字の頻度やモーラの頻度（トークン）に応じて重みの変化を調節するような操作は行っていない。各モーラは、音韻・音読課題の各訓練エポックにおいて、2（例 /re/）から28（例 /ka/）出現した。音読課題では、各文字は各訓練エポックで 1 回だけ出現した。

発達性失読症モデル

このモデルでは、ネットワークが読めるように訓練する前に、音韻表現を破壊することで、発達性失読症をシミュレートした。発達性失読症モデルのシミュレーション方法は Harm and Seidenberg (1999) の方法と同じとした。音韻障害では、文字を読む訓練をする際に、ガウスノイズ ($\sigma = 0.0125$) を重みに注入することで、音韻構成要素内の結合を病変させた。それぞれのモデルでは、20 個のネットワークを構築し、初期化時にランダムな重みを与え、4 万エポックの学習を行った。

結果

ネットワークの応答は、状態が 0.5 以上のすべてのアクティブな音韻ユニットを単純に連結したものである。図 4 は、典型的な健常者のネットワークと典型的な失読症のネットワークの、かな文字と漢字に対する学習過程での性能を示したものである。



40,000 エポックの学習後 20 個の正常な読み上げネットワークは、学習コーパスに含まれる 360 個の仮名および漢字のすべてを正しく発音した。漢字の方がかなよりも急速に性能が向上しており、25,000 エポックまでに漢字は 97.7 %以上、かなは 66.9% の正答率となった。このデータの信頼性を確保するために、全 20 ネットワークの各発達曲線にプロビット解析を適用し、各文字の正答率が 50 %になるまでに必要なエポック数を推定した。その結果、漢字の平均エポック数は 17,398 回、かなの平均エポック数は 22,801 回であった。 $t=38.58$, $df=19$, $p<0.001$ となり、漢字の方がかなよりも有意に早く学習したことがわかった。

一方 40,000 エポックの学習では 20 個のディスレクシアネットワークのうち、360 個の仮名漢字をすべて正しく発音できたものはなかった。40,000 エポックの時点で、漢字の読みの精度 (平均正解率) は 85.71 %, かなの読みの精度は 68.73 %であった。正規モデルと同様に 20 個のディスレクシアネットワークの各発達曲線にプロビット解析を適用し、各スクリプトの正答率が 50 %になるために必要なエポック数を推定した。その結果、漢字の平均エポック数は 19,008 回、かなの平均エポック数は 25,392 回であり、漢字の方がかなよりも有意に早く学習することがわかった ($t=45.20$, $df=19$, $p < 0.001$)。これらの結果は、かな文字に対する漢字の優位性を示すとともに、音韻成分の病変に対して漢字がより頑健であることを示している。

表 1 は 4 万回のエポックにおいて 20 個の健常者ネットワークと 20 個のディスレクシアネットワークの平均サイクル (RT) と 20 個のディスレクシア・ネットワークのひらがな、カタカナ、漢字の発音における平均正解率を示したものである。

Table 1. Mean Cycle (Naming Latencies) and Correct Rate of Normal and Dyslexia Networks in Pronouncing Hiragana, Katakana, and Kanji Characters at 40,000 Epochs.

Model	Cycle/Rate	Hiragana ^a		Katakana ^a		Kanji ^b	
		M	SE	M	SE	M	SE
Normal	Cycle	5.51	0.07	5.51	0.07	4.87	0.02
Dyslexia	Cycle	7.17	0.04	7.17	0.04	7.41	0.06
	Correct rate	0.69	0.02	0.69	0.02	0.86	0.02

^a $n = 71$. ^b $n = 218$.

Table 2. Mean Cycle (Naming Latencies) and Correct Rate of Normal and Dyslexia Networks in Pronouncing Kanji Characters With Three Types of Mora Length at 40,000 Epochs.

Model	Cycle/Rate	Kanji With Single Mora ^a		Kanji With 2 Morae ^b		Kanji With 3 Morae ^c	
		M	SE	M	SE	M	SE
Normal	Cycle	4.68	0.01	4.90	0.02	4.93	0.03
Dyslexia	Cycle	6.69	0.09	7.53	0.06	7.79	0.07
	Correct rate	0.94	0.01	0.86	0.02	0.78	0.03

^a $n = 37$. ^b $n = 145$. ^c $n = 35$.