

知能作用過程のシュミレーション

その1 文字情報の読み取り

寺岡 宏 *矢吹 哲夫

この論文のシリーズは知能作用の過程をパーソナル・コンピュータによってシュミレーションすることを目的とするものである。そのため知能の作用過程を

1. 文字情報の読み取りと意味の照合、意味の推察
2. 外界の情報を受取りこれを統合する神経細胞のニューロンの働き方
3. 意味空間における思考の展開

以上3つの課題に区分して扱うものとする。本論文はその第一段階としての文字情報の読み取りをモデル化しそれらの過程の進行状況をコンピュータ画面に図示することを目的とするものである。さらに各過程に作用する要因の働き具合を比較検討するものである。

文字の視覚的認知の問題としてはアイコン 語の優位性 パターン認識 視覚的記憶などの領域において多くの研究がなされてきた。さらに読語過程の理論的考察としてMorton (1969) による情報収集装置の存在を仮定する Logogen Model、このモデルをさらに数量的に発展させ計算機によるシュミレーションを行なうMcClelland, Rumelhart (1981, 1982) によるInteractive Activation Model, Waltz & Pollack (1985) による自然言語解釈のStrongly Interactive Modelなどが提唱されている。これらの認知モデルはいずれも読語過程の実験心理学によって明らかにされた諸事実を説明するものである。今回のシュミレーションは以上の視点と異なり、パーソナルコンピュータの機能を基礎とし、人間の知能作用という高次な機能にコンピュータの機能をいかに近づけることが可能かを追求することにある。それゆえこの論文は知覚作用のメカニズムを解明するためのものではなく、試験問題の量と妥当性、被験者の疲労 心理的効果の現われ方などの現実的な問題についての解決の方法を示唆するものである。

*北海道大学・理学部・物理学科・大学院博士課程

A : モデルの基本構造

1000Bの情報を持つ文章を乱数を用いて作製し、これを10個の文節に分ける。それゆえ各文章は基本的に100Bの情報量を持つが、文の最後にくる単語の長さの関係上、1つの文の情報量が102Bから103Bになることもある。

文字としては数字とA-Fまでの英字を用いた。このため乱数による数を16進法に転換する事によってそれぞれの文字を作製した。文章の難易度を初級・中級・上級に区別し、初級では用いられる単語の長さを1と2文字、中級では1、2、3文字、上級では1、2、3、4文字とした。文の中でのそれぞれの単語の長さは乱数によって決定された。各クラスの文章の単語の構成は下記の値を示した(10個の文章:10000 B の平均値)。

初級 1語の単語32 2語の単語485

中級 1語の単語11 2語の単語220 3語の単語184

上級 1語の単語8 2語の単語117 3語の単語103 4語の単語113

以上の方法によって作られた上級の文節の一例を下記に記載する。

96 12F 1112 A 2 FC 1192 A 2 A 4 10B0 8 D 1 A 6 1005 8 A 1 D 1
109D 76 FB 11E8 24 154 1075 1 D 1 AA 1116 CC 104 10EE DD 122
1130 91 200 1173 9C 18B. (101 Byte 35 Word)

情報の読み取りを行なう者を学習者とし、そのレベルを初級・中級・上級に区別した。初級者は1、2文字の単語を理解できる能力、中級者は1、2、3文字、上級者は1、2、3、4文字の単語を理解できる能力とした。

各レベルの学習者が記憶として保有している単語の数は情報の理解に重要な要素となる。保有する単語の数が多くなると、読み取られた単語の意味の照合に時間がかかることとなる。このモデルでは最初1、2文字の単語を500個 3文字の単語を1000個 4文字の単語を1000個それぞれ乱数によって作製し、これを学習者が保有する知識とした。

16進法で作製される1、2文字の単語は0からFFまでの256種類である。それゆえ乱数によって0から255迄の数を500個発生させこれを文字化するとき、同一単語が数回登場すると同時に、作られない単語も出てくる。初級の学習者が保有する単語の平均値は219種類(最高228 最低203 標準偏差値5.4)となる。この単語の数に見られる相違は個人の知識内容の相違としての

意味を持つものである。

この関係は中級・上級の学習者の保有する単語にも反映させた。中級・上級のそれぞれ1000個の単語は1から500迄の数を1000回発生させこれを文字化したものである。中級・上級の学習者はともにその対象となる500種類の単語中 平均432個 (最高447 最低414 標準偏差値7.4)を保有する。

学習者の保有する単語(これを記憶とする)はその長さごとに、文字列右端の文字によってそれぞれ16のグループにわけ配列に入力した。

これは学習者の知識の整理をモデル化したものである。

情報の読み取りの操作は、与えられた文章の先頭から1個ずつ文字を取り出し単語を構成し、1つの単語毎にこれを学習者の保有する単語と照合する操作の繰り返しによって進行させた。読み取りの終了した単語を画面に図示し、学習者の保有していない単語を別色で示した。同時に読み取り時間と読み取り情報量の関係をグラフで図示させた。

以上の方法による200Bの文章の読み取り速度は7-8秒を要し、人間の文字情報の読み取り速度に比べ、3-4倍の速さを示す。なお記憶と単語の照合には200Bの文章の場合3-4秒を要し、読み取り速度の約50%が照合の時間に当てられている。1語当たりの照合時間は上級の文章の場合平均68 m sec 初級の文章の場合平均45 m secの値を示した。

情報の読み取り速度を遅らせるため、乱数によって50以上350以下の数(SP)を規定し

FOR I=1 TO SP:NEXT I …… (1)

(1)の操作による待ち時間を1文字毎の読み取りの間に挿入した。SPの平均値=198 SP=1による待ち時間0.45 m sec、その結果1文字毎の読み取りの待ち時間は平均89.1 m secとなる。待ち時間を入れない場合1文字毎の読み取り時間は約40 m secであり、遅延効果により1文字毎の読み取り速度は約130 m secとなり最初に比較して1/3以下になる。

B: 疲労と心理効果によるモデルの修正

文章の読み取りが一定程度継続してゆくとき、疲労によって読み取り速度は低下してゆく。この効果をモデルに追加した。疲労の効果は個人差があり、一定ではないが、情報の読み取り量に対して指数関数的な増加を示すものと

仮定し、次のような一般的な取り扱い方をした。

1. 1文字毎の読み取りの待ち時間を、読み取った単語数に比例して遅くする。
2. 最後の文字の待ち時間を第1文節最初の文字の値(89.1m sec)の2倍とする。

以上の仮定に基づき、SPを次のように修正し修正値(3)を(1)式に用いた。

$$X = ((\text{読み取った単語数}) / (1000\text{Bを構成する全単語数})) * 6.6 \dots\dots (2)$$

$$SP = SP * (10/9)^X \dots\dots\dots (3)$$

1000 Bの文章の場合、(2)式的全単語数は、初級では平均517語 中級では415語 上級では340語となる。以上の待ち時間を用いた各クラスの文章の読み取り時間と情報量の関係を図1、図2、図3に示す。

各カーブは遅延効果によって読み取りが次第に遅くなり、第1文節に比較して第10文節の読み込みが8-9秒遅れる事を示している。単語の照合時間には(1)の遅延効果を作作用させないため第10文節の読み込みは第1文節の1.4-1.6倍の値を示す結果となる。

一定の時間の制約のもとで文章を読む場合、後半になるにつれ、読み取りを速める心理的效果が作用する。この効果として、10文節の最後の読み取りの待ち時間を1文節の最初の値(89.1m sec)と等しくさせることを想定し、

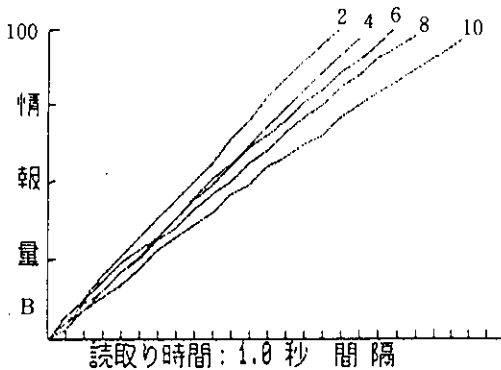


図1. 初級の文節の読み取り速度

知識作用のシュミレーション その1 文字情報の読み取り

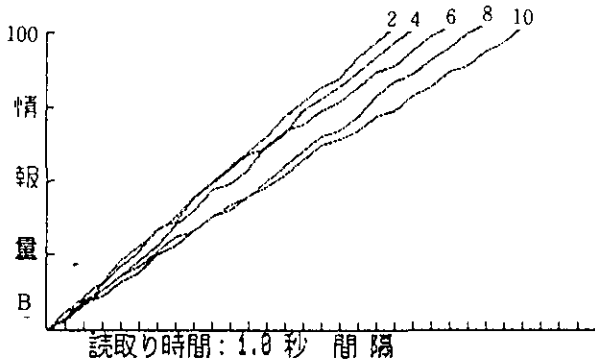


図2 中級の文節の読み取り速度

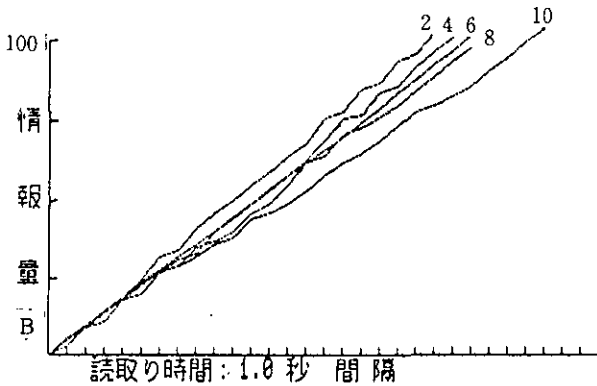


図3 上級の文節の読み取り速度

(3)の値を次の様に修正した。

$$Y = (\text{文節番号} / 2) + 1 \dots\dots\dots (4)$$

$$SP = SP - SP * (0.3) * (Y / 5.7)^4 \dots\dots\dots (5)$$

(5)に用いるSPは(3)式によるものである。以上によるSPの修正値を用いて(1)

を実行した。第一文節の最初の読み取りの待ち時間を1とした場合の変化を
図4に示す。

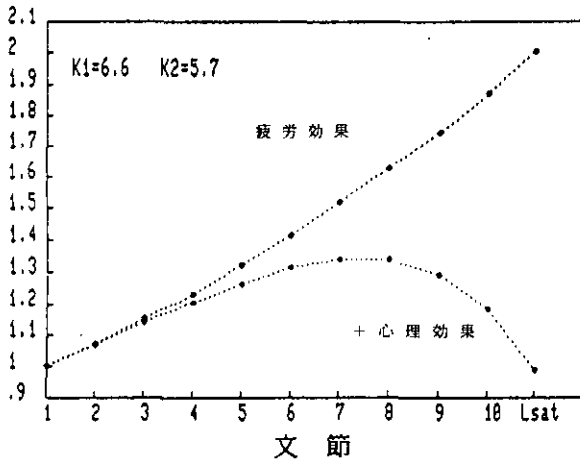


図4. 疲労効果と心理効果の現われ方

(縦軸の値は第一文節最初の文字の読み取り速度を1とした相対値)

以上の仮定に基づく遅延効果と心理的効果は個人の能力・素質・与えられた状況によって多様に変化する。それゆえ(2)から(5)の式のパラメータは個人の特徴に即して変化させる事が可能であり、そのことによって情報の読み取り過程を解析する糸口が見出される。(3)式のXの代わりに X^2 , X^3 ・・・を用いる時 遅延効果の働き方はより指数関係的な特徴を示す。図5は X' を用いた場合の結果と心理効果を示す。

(5)式の(0.3)は心理効果の強さを平行移動的に調整するパラメータ、(Y/5.7)の5.7の値は心理効果の現われる文章の場所に関するパラメータでこの値が小さい程、早く心理効果が発現する。累乗を示す4の値は心理効果の現われる場所と、最後の文字の読み取り待ち時間を幾らにするかによってその大きさを変化させる。図5の心理効果は X' を用いた場合の

$$SP = SP - SP * (0.15) * (Y/5.9)^{12}$$

による値を示す。

知能作用過程のシュミレーション その1 文字情報の読み取り

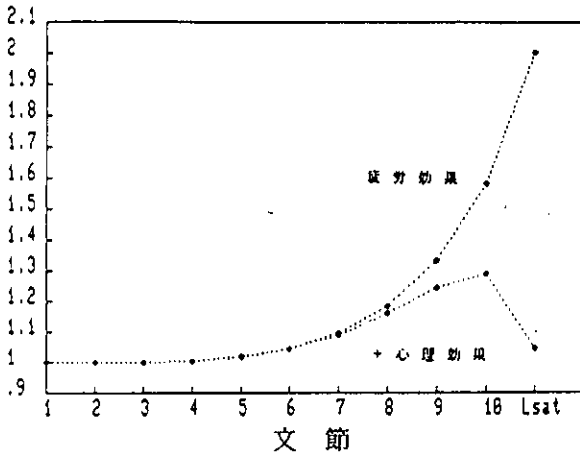


図5. 疲労効果と心理効果の現われ方

(縦軸の値は第一文節最初の文字の読み取り速度を1とした相対値)

(5)によるSPの値を用いた場合の各クラスの文章の読み取り時間と情報量の関係を図6、図7、図8に示す。各クラスとも心理効果によって最後の文章の読み取り速度が最初の文の読み取り速度に近い値をしめしている。

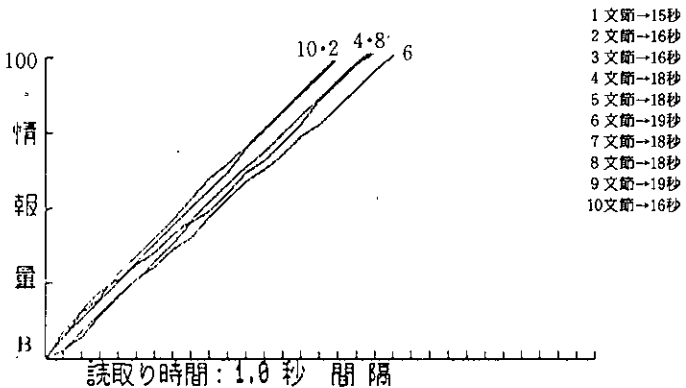


図6. 心理効果を加えた初級文節の読み取り速度

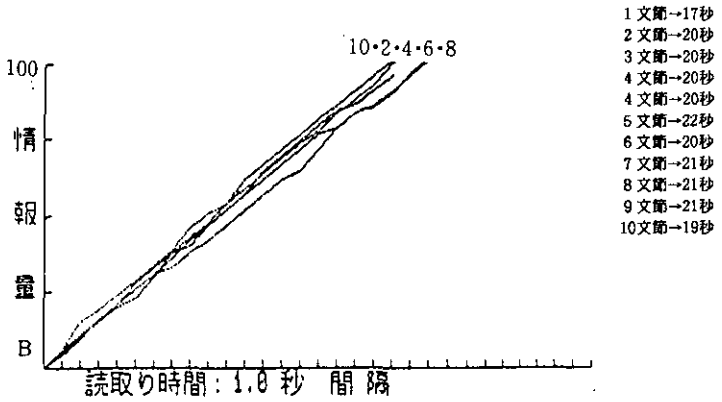


図7. 心理効果を加えた中級文節の読み取り速度

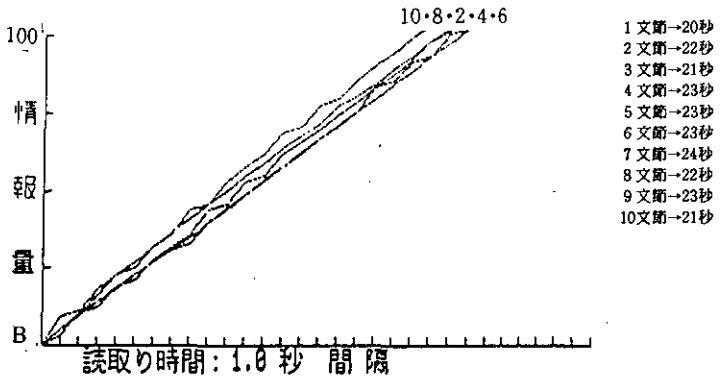


図8. 心理効果を加えた上級文節の読み取り速度

C : 読み取りのエラーと意味の不明な単語に対する推察

文字を読み取る過程において読み取りのエラーを伴うことがある。この点を次の仮定によってモデルの中に組み入れた。

読み取りエラーの発生する可能性はある一定の確率において潜在的に存在するが、これがエラーとして実際に発生するか否かはその時の文字の読み取り速度と継続時間（疲労）によって影響される。この考えを次の式によって示した。

$$60+(J-1)*10>SP \text{ AND } \text{INT}(1+\text{RND}(1)*6)>5 \dots\dots\dots (6)$$

(6)における Jは読み取りの継続状況を示すファクターとしての文節の番号、SPは(5)における値を用いる。読み取りの1文字毎に(6)式を計算し、論理式の真が成立するときはその文字の読み取りのにエラーが発生するものとした。読み取られた文字を表示するときエラーの文字は*によって示した。

学習者がある単語を読み取った場合、学習者にその単語についての知識がなく、意味の不明な場合が起こる。この場合学習者は今まで継続してきた思考の流れを一旦中断し ある一定時間後にその語の意味の推察を行なう。この思考の中断時間として FOR T=1 TO SS:NEXT T を仮定し、初級の単語の場合 SS=500, 中級 SS=900, 上級 SS=1300の値を用いた。

意味の不明語の推察はその前後の単語の意味によってなされる。この点を次の仮定を用いてモデルに加えた。該当する単語の前後の単語を選びこれをWA WBとする。WA WBの各文字列の右側から文字を2個取り出し、これをWA1,WA2, WB1,WB2とする。この場合文字列の右端をWA1, WB1 右端から2番目の文字をWA2,WB2とする。WA1, WA2, WB1, WB2を数値化しWA1, WA2の平均値をWA3, WB1,WB2の平均値をWB3とする。意味の不明語についても文字列の右端から文字を2個取り出し、これをWX1, WX2とし平均値をWX3とする。意味の不明語の文字列の長さをLとする。WA3, WB3のうち値の大きい方をW1とし、小さい方をW2とする。下記の(7)の条件が成立する時は意味の不明語について、意味の推察が可能なケースとした。

$$W2-(3-L)<WX3 <W1+(3-L) \dots\dots\dots (7)$$

(7)における(3-L)の項は単語の意味の推察における難易度を単語の長さによって調整したものである。単語が長くなる時意味の推察を可能とする幅は狭くなる。

読み取りエラーの結果 文字の代わりに、*を含む単語は全て意味不明語

となる。*が文字列の右側2列に含まれない場合は(7)による意味の推察の対象とした。WX2が*のときは $WX2=0$ $WX3=WX1$ として(7)の判定を行なった。WX1が*の場合は推察不可能なケースとして扱った。単語の意味の決定に関し、文字列の位置に違いを仮定し、右端の文字の読み違いを決定的なエラーとし、修復不可能なものとする仮定に基づくものである。

以上のモデルによる文章の読み取り過程におけるエラー（*印の文字）、意味不明語（アンダーライン）、意味の推察可能語（2重ライン）の発生状況について上級クラスの文節について例示する。

例1：上級第6文節

C7 CE 10EA 1C 1A5 10EF E2 288 104F 77 CD 1041 5D 25E
 11D1 DB *66 1084 8E 19A 107A CB 120 10D9 14 18E 1126 91 11A
*1D0 99 167 10AE D8 1E0

例2：上級第8文節

67 1C6 1100 84 221 1*6D 0E 1FF 1*37 4 20F 1071 B4 E
 6 1020 85 1D5 1175 E5 239 10C4 89 DC 1022 10 29E 1038 96 118
 109B A6 1E6 1145 8B *8E

例3：上級第10文節

1 257 105D E4 263 106A 03 23D 1096 49 23C 10FB 48 F0 11B 1
 E7 1FD 11AA DE 1C1 104D 3 10DE B* 27* 11CA D* CD 1040
 D4 204 1167 D8 240 11A8

D：『Zipfの法則』に基づく単語の 使用頻度の修正と保有知識の作成

以上のモデルにおいて読み込みを行なう文字と、学習者の知識は乱数によって無作為に作成したものであった。それゆえ1個の文節において同じ単語が繰り返して用いられることは殆ど有り得ない。しかし実際の文章においては単語の使用頻度には(8)の関係が見られ（Condon 1928）、これはZipfの法則（Zipf 1949）に従うものとされている。

$$\text{使用頻度}X\text{位の単語の出現頻度} = k / X \dots\dots\dots (8)$$

k : 各単語の出現頻度の総和を1とするための係数

以上の関係をモデルに取り入れるため、次の方法によって単語の母集団を作成した。初級レベルの学習者の単語の母集団として500個の単語を含むサンプルを作成した。各単語は乱数によって1000以下の数を発生させ、これを16進法に転換して文字化した。1個ずつの単語に使用頻度の順位(X)と(8)式による出現頻度を与えた。

中級レベルの単語の母集団として750個、上級レベルの単語の母集団として1000個の単語を含むサンプルを作成した。

中級レベルの単語の母集団において、出現頻度の1位から3位までの単語は初級レベルの単語と同じものを用いた。4位以下750位までの単語は1000以上4095までの数を乱数によって発生させ、これを16進法に転換して文字化した。同時に各順位毎に $P = \text{INT}(\text{RND}(1) * 3 + 1)$ をもとめ $P = 1$ の場合は初級の母集団の4位以下の単語を順に中級集団の単語として取り入れた。その結果中級の母集団750個の単語の中には、初級と同じ単語が約250個含まれている。上級レベルの単語の母集団において出現頻度の1位から4位までの単語は中級レベルの単語と同じものを用いた。5位以下1000位までの単語は4096以上60000までの数を乱数によって発生させ、これを16進法に転換して文字化した。中級の場合と同様1/3の確率において、中級の母集団の5位以下の単語を順に上級集団の単語として取り入れた。その結果上級の母集団1000個の単語の中には、中級の母集団を構成する単語が上位から約330個含まれている。使用頻度X位の単語に(8)式による出現頻度を与えた。

表1に各レベルの上位20位迄の単語と出現頻度(%)を示す。

乱数によって(8)式の出現頻度に相当する値を作り、その値によって各レベルの母集団から相等する頻度の単語を選出し文章を作成した。この方法による文章では出現頻度の高い単語の場合、同じ単語が2回続けて並ぶケースや、1個おきに同じ単語が並ぶケースが見られる。現実の文で以上の様なケースは殆ど見られない。そのため文章の流れの中で単語4個の配列中に同一単語が現われないように単語の配列に修正を加えた。これはZipfの法則の効果を弱めることになる。(修正の有無による文章の読み取り効果の比較は後述)

表1. 初級、中級、上級の単語の母集団における
頻度20位までの単語

	初 級		中 級		上 級	
	単語	出現%	単語	出現%	単語	出現%
1位	24A	14.7214	24A	13.8923	24A	13.3592
2位	215	7.3607	215	6.9465	215	6.6796
3位	49	4.9072	49	4.6309	49	4.4531
4位	30B	3.6803	A5C	3.4732	A5C	3.3396
5位	EF	2.9443	30B	2.7786	8D17	2.6718
6位	1C1	2.4536	6CC	2.3155	30B	2.2266
7位	EE	2.103	6CB	1.9847	480D	1.9084
8位	1FC	1.8402	47A	1.7366	4803	1.6699
9位	2F	1.6357	79A	1.5436	1B1E	1.4844
10位	23F	1.47221	A8D	1.3893	579E	1.3359
11位	131	1.3383	EF	1.263	90C7	1.2145
12位	207	1.2268	CC3	1.1577	6CC	1.1133
13位	225	1.1324	C99	1.06871	BBA7	1.0276
14位	108	1.0515	1C1	.992294	B876	.964199
15位	3D8	.9815	A0E	.926202	6CB	.890598
16位	2DC	.919998	856	.868301	8730	.835001
17位	2CE	.866	EE	.817297	65E0	.7858
18位	124	.817896	B12	.771802	47A	.742203
19位	3D0	.774806	F2D	.731198	9ADA	.703101
20位	2B8	.736096	DB3	.694605	EA67	.667996

文 字 列

E50 (67)	90C7 (11)	83C1 (166)	13FC (103)	2878 (346)	F579E (10)	6CB (15)	49 (3)
EA67 (20)	F452 (912)	F58 (179)	918C(162)	A628 (29)	579E (10)	79A (23)	F414 (99)
645 (83)	DB3 (58)	3A0F (41)	59B4 (223)	1B1E (9)	14E (180)	30B (6)	9FA (60)
EA67 (20)	BBA7 (13)	30B (6)	24A (1)	87F (62)			
215 (2)	344F(143)	24A (1)	6510 (121)	9FA (60)	58D2(171)	4803 (8)	A 6150 (121)
BBA7 (13)	A8D (28)	A5C (4)	65E0 (17)	B390 (26)	D25 (385)	A41 (98)	F90 (785)
6B27(642)	24A (1)	F87A(375)	1036 (263)	8D17 (5)	9BF0 (95)	9F39 (57)	9B70 (184)
B876 (14)	24A (1)	215 (2)	EA67 (20)				

下にこの方法によって作られた上級の文節を2個例示す。()内の数字は使用頻度の順位 (X) を示すものである。

学習者の保有する単語も、各レベルにおける使用頻度の高い単語ほど保有する確率は高くなる。学習者が初級の場合、初級単語の母集団から200個の単語を保有する者と300個の単語を保有する者を作った。

単語の取り出しはZipfの法則に従う頻度を乱数によって作成し、母集団からその値に相当する単語を抽出した。同じ単語が繰り返し母集団から取り出されるため、単語の種類がそれぞれの数になるまで抽出を繰り返した。それゆえ使用頻度の高い単語ほど母集

団から取り出される可能性は高くなる。

中級の学習者の場合、単語の母集団から250個の単語を保有する者と400個の単語を保有する者を作った。上級の学習者の場合、上級単語の母集団から300個 500個 700個の単語を保有する3通りの者を作った。

以上の方法による学習者の保有する知識の一例として、上級500語を保有するケースでは、順位200位迄の語を186個、順位201-400位迄の語を122個、順位401-600位迄の語を88個、順位601-800位迄の語を54個、順位801-1000位迄の語を50個もつことが観察された。

以上のモデルによる上級の3種類の学習者について、文字情報の読み取り結果を示す。

表2. Zipfの法則に基づく上級の文章の読み取り(1000B 294語)

学習者の 単語数	読み取り時間 秒		意味の不明語数		意味不明語%	
	無修正文	修正文	無修正文	修正文	無修正文	修正文
300	21.4	22	61	82	20.7	28.2
500	18.4	19.7	19	42	6.4	14
700	17.6	19	10	26	3.4	8.6

表2の結果が示すように、上級の単語1000語の母集団の内700語を保有する場合、判定の出来ない単語の出現率は修正文で8.6% 無修正文で3.4%を示す。この値は学習者の保有しない300個の単語の出現率を示すものであり、無作為的に出現する場合には30%の値を示す。表2に見られるこの値の減少はZipfの法則による効果を示すものである。また同一単語の繰り返しを避けた修正がZipfの法則の効果を減少させていることが見られる。

同様の関係は上級の単語保有500語、300語のケースでも観察され、また保有単語数が多い程、文章の修正によってZipfの法則の効果は減少することが認められる。

単語の使用頻度と思考に用いられる知識の使用頻度とは必ずしも同一とは考えられないが、しかし両者の間には高い類似性が考えられる。この類似性に基づくとき必要とされる知識量の70%を保有することによって現実の問題に対しては90%以上の理解が可能とされること この学習の効果は問題文がZipfの法則により正確に従う程強く現われること、学習の効果は使用頻度の高い知識を効果的に保有すること すなわち基礎をしっかりと把握することの重要性などが今回のシュミレーションによって示された結果である。

E : まとめ

以上のモデルによって文字読み取り過程における速度 疲労 心理効果 読み取りエラー 意味の照合と推察 読み取り量と時間との関係 保有知識の数と学習効果との関係などをシュミレートすることが可能である。表3はこのモデルによるZipfの法則に基づいて作られた1000 Bの文章の読み取り過程にみられるエラー、意味不明語、意味の推察可能語の発生結果を文節毎に集計したものである。

表3-A. 文節毎の読み取り状況

文のレベル 上級 学習者レベル 上級 300語

文 節	時間(秒)	Byte数	意味の判定の出来ない単語数			意味・推察 の出来た語
			読み取り エラー	非保有語	%	
1	19	102	1	7	24.1	2
2	24	104	2	13	44.8	5
3	19	103	1	6	20.6	1
4	20	103	3	6	21.4	1
5	23	102	1	8	28.5	2
6	23	103	3	8	28.5	0
7	24	105	0	9	30	1
8	23	103	3	8	27.5	2
9	21	102	0	5	16.6	0
10	24	105	5	12	40	0
平均値	22	103.2	1.9	8.2	28.2	1.4

このモデルを個人の読み取りのシュミレーションとして用いるためには計算式における各パラメータに個人の特性を反映させる数値を用いることが必要である。一方学習者を例えば読み取り速度は速いが疲れ易く、時間を強く意識するタイプ、速度は遅いが疲れ難く、時間による心理的圧迫を殆ど感ぜることのないタイプなど幾つかの類型に区別をするとき、それらの特徴を各パラメータに付与し、読み取り過程の比較をすることは可能である。

以上のような類型化によるシュミレーションは、試験の問題量 試験時間の長さ、一定の正確さを保つための休憩時間のとり方など、いくつかの課題の解決に有効な方法となるであろう。また問題文を作成する時 Zipfの法則

知能作用過程のシュミレーション その1 文字情報の読み取り

表3-B. 文節の読み取り状況
文のレベル 上級 学習者レベル 上級 500語

文 節	時間(秒)	Byte数	意味の判定出来ない単語数			意味・推察 の出来た語
			読み取り エラー	非保有語	%	
1	19	104	1	6	20	2
2	18	105	0	3	10	1
3	17	102	1	2	6.6	1
4	18	103	0	3	10.3	0
5	20	102	1	5	16.1	1
6	23	104	1	6	20	1
7	23	104	7	7	22.5	2
8	19	103	0	1	3.2	0
9	20	102	2	4	13.7	1
10	20	102	3	5	17.2	0
平均値	19.7	103.1	1.6	4.2	14	9

表3-C. 文節毎の読み取り状況
文のレベル 上級 学習者レベル 上級 700語

文 節	時間(秒)	Byte数	意味の判定出来ない単語数			意味・推察 の出来た語
			読み取り エラー	非保有語	%	
1	18	104	1	3	10	2
2	19	105	1	3	10	0
3	17	102	0	0	0	0
4	17	103	0	1	3.4	0
5	21	102	3	6	19.3	1
6	20	104	2	3	10	0
7	23	104	6	6	19.3	2
8	19	103	0	0	0	0
9	18	102	0	1	3.4	0
10	18	102	3	3	10.3	0
平均値	19	103.1	1.6	2.6	8.6	5

に基づく第何位程度の知識を用いるかは、問題の難易度の定量化を示すものとなる。同時にその問題に対する学習者の理解度から学習者が有する知識量を定量的に推察することが可能とされる。また問題中の知識の使用頻度をZipfの法則の値と比較することも問題の妥当性をはかる一つの指標として用いられるであろう。

引用文献

- (1) Morton, J. 1969 Interaction of information in word recognition. *Psychological Review*, 76, 165-178.
- (2) McClelland, J.L. & Rumelhart, D.E. 1981 An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88, 375-407.
- (3) Rumelhart, D.E. & McClelland, J.L. 1982 An interactive activation model of context effects in letter perception: Part 2. *Psychological Review*, 89, 60-94.
- (4) Waltz, D.L. & Pollack, J.B. 1985 Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, 9, 51-74.
- (5) G.K. Zipf 1949 Human behavior and the principle of the least effort. Addison, Wesley Press Inc.

Simulations of Operating-processes of the Human Intellect

Part 1. Reading of Literal Information

Hiroshi Teraoka and Tetuo Yabuki

This paper constitutes a model which simulates the processes of the reading of 1000 Bytes of literal information. These variable parameters are contained in the present model: exhaustion by reading, mental pressure, reading error, and guessing about unknown words.

Readings were carried on each letter of a word. When the computer recognized a word, that word was checked against the memory of the person tested. Reading speed was controlled by the waiting time that was inserted between each letter reading.

Word-frequency principles based on Zipf's law were applied to make the test-sentence by computer. The contents of knowledge of persons tested were produced according to the word-frequency principles based on Zipf's law.

These simulations indicated characteristic features existed in the reading processes, such as relations found between reading speed and accuracy; mental pressure and delay by exhaustion; and results of the test and knowledge-volume in the person tested.