

# 日本文の可読性の測度と表示速度への応用†

浅野 陽子<sup>††</sup> 小川 克彦<sup>††</sup>

本論文では、学年レベルを基準とした新たな日本文の可読性の測度 (RGV; Reading Grade Value と呼ぶ) を定義する。さらに、RGV の 1 つの応用として、コンピュータディスプレイに表示する文書のスクロール速度を、RGV の値に応じて設定する方式を提案し評価する。まず、小学校から高等学校までの国語の教科書に掲載された文書中の、文字種や句読点などの字面要因の統計量を対象に因子分析を行い、高い寄与率を示す要因を選択する。これらの要因を説明変数として、学年レベルを目的変数とする重回帰分析を行い、その結果得られたひらがな、カタカナ、句点の 3 つの要因を変数とする線形式により、RGV を定義する。次に、文書の RGV の値と、ディスプレイ上に表示した文書を人が読む際の黙読速度に、相関があることを実験により示す。この性質を利用して、文書の RGV 値に応じてスクロール表示速度を設定する方法を述べる。さらに、表示速度が一定の方式との比較実験結果から、RGV により速度を設定する方式が、読み手の黙読速度に、より適合していることを示し、RGV の有用性を確認する。

## 1. はじめに

英文の可読性の測度に関する研究は、1940 年代から多く行われており<sup>1)</sup>、特に、Flesch<sup>2)</sup> や Dale-Chall<sup>3)</sup> の測度は広く用いられてきた。これらは、文や単語の長さ、語彙の頻度に関する 2 つの変数の線形式で算出され、小学生から高校生を対象とした各学年の読解力テストの結果を基準としている。

これらの可読性の測度は、子供向けの本<sup>2)</sup>、新聞<sup>3)</sup>、さらには海軍での訓練教本<sup>4)</sup> などを書く際の、読み手のレベルに合わせた文書を書く基準として用いられている。また、読み手の学歴と文書の学年レベルとの関係から、読み手のリーディングスキルを予測するのも用いられており<sup>5)</sup>、その有用性を示す実例は多い。

これに対し、日本文は、英文のように単語ごとにかち書きされず、べた書きであり、また、漢字、ひらがな、カタカナ、英字といったいろいろな文字の種類が用いられる。そのため、英文の可読性の測度を、そのまま日本文に適用することはできない。しかし、その考え方は、日本文の可読性の測度に反映されている。

日本文の可読性に関しては、1960 年代から、語彙の難易度や文の長さによる査定基準などの研究が行われている<sup>6)~9)</sup>。さらに矢頭は、一文の平均文字数と 100 字当たりの漢字とカタカナの頻度を変数とし、読みやすさの主観評価を外的基準とする線形式による測度を提案した<sup>10)</sup>。また、建石らは、対象とする文書の因子

分析を行い、その因子から求めた標準偏差と平均値を基準として、文書中の 6 つの字面要因を変数とする線形式で可読性の測度を定義した<sup>11)</sup>。

そこで、本論文では、従来測度の作成法を基本とし、新たに客観的な基準として学年レベルを採用した日本文の可読性の一測度 (RGV; Reading Grade Value と呼ぶ) を提案する。さらに、RGV の 1 つの応用として、コンピュータディスプレイに表示する文書のスクロール速度を、RGV の値に応じて設定する新たな方式を評価し、RGV の有用性を確認する。

まず、小学校から高等学校までの国語の教科書に掲載された文書を対象に、ひらがな、カタカナ、漢字、句読点などの 14 種類の字面要因の統計量を算出して因子分析を行い、高い寄与率を有する要因を選択する。そして、これらの要因を説明変数とし、学年レベルを目的変数とする重回帰分析を行って、ひらがな、カタカナ、句点の 3 つの要因を変数とする線形式により RGV を定義する。

さて、この RGV は、英文と同様、読み手のレベルに適した文書を書く基準として、また、文書を読む基準としても利用することができる。そこで、まずディスプレイ上に表示した文書を人が読む際の黙読速度と、その文書の RGV 値とに相関があることを実証する。この性質を利用して、文書の RGV 値に応じてスクロール表示速度を設定する方法を述べ、さらに表示速度が一定の方式との比較実験により、読み手の黙読速度との適合性を評価する。

## 2. 日本文の可読性の測度

国語の教科書の文書を対象に、文中のひらがな、カ

† A Readability Measure for Japanese Texts and Its Application to Presentation Rate by YOKO ASANO and KATSUHIKO OGAWA (NTT Human Interface Laboratories).

†† NTT ヒューマンインタフェース研究所

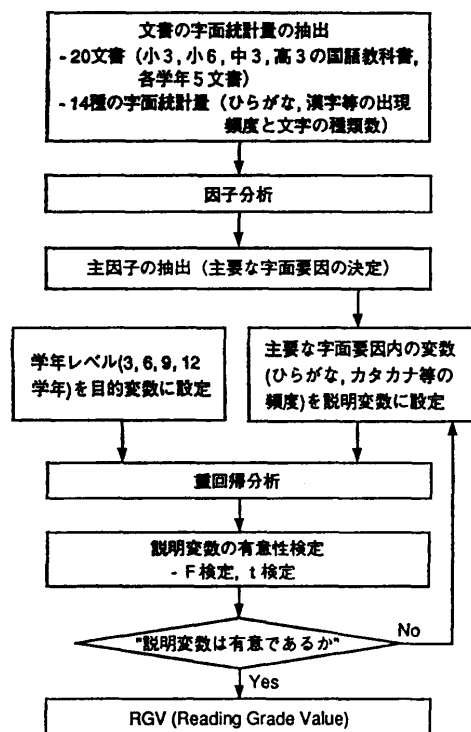


図1 教科書の学年レベルを基準とした日本文の可読性の測定 RGV を定義する手順

Fig. 1 Procedure to determine RGV (Reading Grade Value) for the readability measure for Japanese text.

タカナ, 漢字, 句読点などの字面要因の因子分析を行い, 高い寄与率を示す因子を抽出する。これらの因子の重回帰分析を行い, 日本文の可読性の測定 (RGV) を定義する。図1にこれらの手続きを示す。さらに, 英文の可読性の測定が広く活用されていることから, RGV とこの英文の測定とを比較考察し, RGV の応用分野を考える。

## 2.1 字面要因の因子分析

因子分析の対象とする文書は, その違いが比較的明確になるように, 3 学年おきの国語の教科書から選択することとし, 小3, 小6, 中3, 高3の国語の教科書から各5編, 計 20 編を選択した (付録参照)。なお, 各教科書での文書は, 以下のとおりである。

- 散文のみ (箇条書き, 韻文等を含まない)
- 会話文があまりない
- 物語文は避ける (著者の作風による影響が大きい)
- 図を用いた説明がない

これらの文書中に出現する 6 文字種 (ひらがな, カタカナ, 漢字, 英字, 数字, 記号) 各々の全文字数に

対する出現頻度の割合 (各文字種の文字数/全文字数, %) と文字の種類数 (各文字種ごとの出現種類数), さらには, 文の区切りに関する統計量として, 句点の全文字数に対する出現頻度の割合 (%) と 1 文当たりの読点の出現頻度の割合 (%) の, 計 14 種類の字面要因を選択した。なお, 「」; % ( ) は記号に分類した。

20 の文書から求めた 14 の字面統計量に対し, 主因子法<sup>12)</sup>を適用し, さらに varimax 直交変換を行って, 各因子に寄与する変数が独立となるような因子を求めた。表 1 は, 高い寄与率を示す順に 4 つの因子をあげ, それぞれの各字面要因の因子負荷量を示している。各因子で, 因子負荷量の絶対値が 0.8 以上の変数を負荷の大きい順に挙げると, 次のようになる。

第 1 因子…ひらがな頻度, 記号種類, 記号頻度

第 2 因子…数字頻度, 数字種類, 英字頻度, 英字種類

第 3 因子…カタカナ頻度

第 4 因子…句点頻度, 読点/句点

これから, 第 1 因子はひらがなと記号, 第 2 因子は数字と英字, 第 3 因子はカタカナ, 第 4 因子は文の区切りに関する因子ということがわかる。ここで, 記号, 英字は, 出現頻度が全体の 1 % 未満であり, ひ

表 1 教科書文書に含まれる字面要因の因子負荷量  
Table 1 The results of factor analysis for 14 statistical data for characters in texts.

| 字面要因   | 第1因子  | 第2因子  | 第3因子  | 第4因子  |
|--------|-------|-------|-------|-------|
| ひらがな頻度 | -.877 | -.181 | .265  | -.17  |
| ひらがな種類 | .209  | -.041 | .728  | .455  |
| 記号頻度   | .801  | .148  | .17   | -.207 |
| 記号種類   | .837  | .118  | .168  | .106  |
| 数字頻度   | .023  | .976  | .028  | .042  |
| 数字種類   | .021  | .974  | .031  | .041  |
| 英字頻度   | .151  | .961  | .069  | .036  |
| 英字種類   | .257  | .852  | .137  | -.03  |
| カタカナ頻度 | .131  | -.161 | -.951 | .021  |
| カタカナ種類 | .706  | -.126 | -.342 | .044  |
| 読点/句点  | -.141 | -.075 | .058  | .916  |
| 句点頻度   | -.361 | -.187 | -.172 | -.856 |
| 漢字頻度   | .732  | .27   | .424  | .181  |
| 漢字種類   | .612  | .113  | .529  | .454  |

注: 各因子を特徴付ける字面要因が対角線上に並ぶように配置した。

らがなの平均出現頻度 62%, カタカナの平均出現頻度 6% に比べて極めて小さい。そのため、各因子を代表する、すなわち文書の特徴付ける字面要因としては、ひらがな、数字、カタカナ、文の区切りに関する要因が適切と考えられる。そこで、統計量の統一性を考慮し、ひらがな、数字、カタカナ、句点の頻度の 4 つの字面要因を、次のステップである重回帰分析の対象とした。

## 2.2 RGV の定義

前節の分析に基づき、ひらがな頻度、数字頻度、カタカナ頻度、句点頻度の 4 変数を説明変数とし、各々の文書の学年レベルを目的変数として重回帰分析を行った。学年レベルとしては、小学 3 年レベルを 3、小学 6 年レベルを 6、中学 3 年レベルを 9、高校 3 年レベルを 12 として用いた。その結果、重相関係数  $R=0.879$ , 自由度調整済み寄与率  $R^2=0.712$  の線形式を得た。

しかし、上記 4 変数各々に対する  $t$  検定の結果は、数字頻度に関しては、 $t=1.296 < t(19, 0.05)$  となり、有意な結果が得られなかった。これは、数字の頻度が、他の 3 つの文字種の頻度に比べ、極めて少ないためと考えられる。

そこで、数字頻度を除いた 3 変数に対して同様に重回帰分析を行った結果、 $R=0.865$ ,  $R^2=0.7$  となり、各変数に対する  $t$  検定の結果もすべて 95% 有意となった。なお、 $t$  検定の有意水準が次に低いひらがな頻度を除いた 2 変数に対して重回帰分析を行った結果は、 $R=0.813$ ,  $R^2=0.62$  となり、3 変数に比べて低くなってしまった。

以上の解析結果より、RGV をひらがな頻度、カタカナ頻度、句点頻度の 3 変数を用いた次式で定義する。

$$RGV = -0.17ph - 0.28pk - 3.49pe + 27.62 \quad (1)$$

RGV : 学年レベルを基準とした可読性の測定  
 $ph$  : ひらがなの出現頻度 (%)  
 $pk$  : カタカナの出現頻度 (%)  
 $pe$  : 句点の出現頻度 (%)

(1)式では、ひらがな、カタカナが多く、句点が多くて一文の長さの平均が短い文書ほど、学年レベルに応じた RGV の値が低い。一方、漢字はひらがなと負の相関が高いため、漢字の頻度が高くなるとひらがなとカタカナの頻度が低くなる。すなわち、ひらがな、カタカナが少なく、漢字が多いほど RGV が高いと

いえる。

新聞、雑誌などの一般の文書の RGV を求めることにより、その文書がどの学年レベルに対応するかがわかる。雑誌の RGV の値は 7~13、新聞は 9~16、新書は 7~14 と、大部分の一般文書は 7~16 に分布している<sup>13)</sup>。小学校から高等学校までの学年レベルは 1~12 の範囲であり、RGV の値が 13 以上の学年レベルはない。もちろん、13 以上の RGV の値は、教科書の学年レベルを外挿した値としてとらえることもできる。しかし、この測定は、文書の学年レベルを定めることが主たる目的ではなく、学年レベルを基準として、字面要因の統計量から文書を定量的に測り、他の文書との相対比較を行うための基準としての効用を狙いとしている。

## 2.3 英文と日本文の可読性の測定の比較

RGV の応用分野を探るため、広く活用されている代表的な英文の可読性の測定 (Flesch および Dale-Chall の測定) と RGV との定義式の比較を行い、その相違を考察する。

Flesch の測定は次式で定義される<sup>2)</sup>。

$$R.E. = 206.835 - 0.846wl - 1.015sl \quad (2)$$

$R.E.$  : reading ease (0~100)

$wl$  : 100 語当たりの音節数

$sl$  : 1 文当たりの平均単語数

$wl$  の値が大きくなり、音節数が多く長い単語ほど難しい語が多いと考えられる。また、 $sl$  の値が大きくなり、一文当たりの単語数が多く長い文書ほど、複雑な構文になると考えられる。したがって、Flesch の測定では、 $sl$ ,  $wl$  が大きい文書ほど、 $R.E.$  の値が小さく評価される。

一方、Dale-Chall の測定は次式で定義される<sup>3)</sup>。

$$X_{c50} = 0.1579x1 + 0.0496x2 + 3.6365 \quad (3)$$

$X_{c50}$  : 読解力の標準テストで正答率 50% となる学年のレベル (5~10)

$x1$  : 基本単語リストに載っていない語の割合 (%)

$x2$  : 1 文当たりの平均単語数

$x1$  の値が大きいくほど、普段あまり使われない馴染みのうすい単語が多く使われていることになる。また、Flesch の測定の  $sl$  と同様に、 $x2$  の値が大きくなり、一文当たりの単語数が多く長い文書ほど、複雑な構文になると考えられる。したがって、Dale-Chall の測定では、 $x1$ ,  $x2$  が大きい文書ほど、 $X_{c50}$  の値が大きく評価される。

さて、RGV の変数で考えると、 $ph$ ,  $pk$  が小さいほどひらがなやカタカナが少なく、漢字が多くなるため、一文字当たりの音節数が多くなる。さらに、ひらがなで表現した単語（例：なく）より、漢字で表現される単語（例：（人が）泣く、（鳥が）鳴く）は特定の意味を有することが多いため、 $ph$ ,  $pk$  は、 $x1$  と同様に単語の平易性の意味を有すると考えられる。また、 $pe$  が小さくなるほど、全文字数に対する句点の頻度が少なくなり、1文の長さが長くなるので、 $pe$  は、 $sl$ ,  $x2$  と同様の意味を有すると考えられる。

以上の考察から、日本文と英文の各測度で使用される変数は各々異なるが、変数の表現する意味は同等となる。すなわち、言語は異なっても、文書の可読性を表す測度は、同等の構造を有するといえる。

このため、英文の可読性の測度と同様に RGV も、新聞や雑誌、あるいはテキストやマニュアルなどを書く際の、読み手のレベルに合わせた文書を書く基準として、さらには、文書を読む基準として有用であると考えられる。文書は、紙面上で読むだけでなく、電子メール、パソコン通信、データベース、ビデオテックスなど、コンピュータディスプレイ上の文書を読む機会も増加している。そこで、読み手の黙読速度に合わせて、文書のディスプレイ表示速度を設定する基準として、RGV を応用することが考えられる。

次章以降では、この RGV の1つの応用について検討し、RGV の有用性を確認する。まず3章で、文書の黙読速度と RGV との相関関係を調べる。もし、相関が得られれば、文書の RGV 値を基準としてディスプレイへの表示速度を設定することにより、ある文書の実際の黙読速度がわからなくとも、その文書の RGV 値から黙読速度を予測し、それを基準として読み手の黙読速度に適したディスプレイへの表示速度を求めることができる。この RGV を基準とした表示速度の設定方法は4章で述べる。なお、ディスプレイへの表示方法には、ページ単位のブロック表示なども考えられる<sup>14)</sup>が、本論文では、現状で多く用いられているスクロール表示に着目して検討を進めることとした。

### 3. 黙読時間と RGV との関係

ここでは、人がディスプレイ上にスクロール表示された文書を黙読する実験を行い、測定した黙読時間とその文書の RGV との相関関係を考察する。

#### 3.1 黙読実験の方法

(1) 被験者：被験者は、10名（女性6名、男性4

名）であり、いずれもコンピュータのディスプレイを見ながらキーボードを使用することに習熟している。

(2) 文書：実験に用いた文書は、RGV 値が幅広く分布するように考慮し、因子分析に用いた小3、小6、中3、高3の国語の教科書から各1編、さらに論文、特許各1編、計6編とした。文書の長さは1,100～2,300字であり、RGV 値は3.2～18.6であった。

(3) 実験装置：市販の16bit パソコンを使用した。文書は、14インチディスプレイ上に40字×12行のサイズで、ネガティブ表示とした。

(4) 実験課題：被験者には、ただ単に「黙読してください」と指示し、ディスプレイに表示された文書を黙読してもらった。実験終了後、各文書の読み方についてインタビューした。文書は、画面の下段から1行単位で表示され、被験者は、自分の読みのペースに合った速度で文書を読むことができるように、被験者のキー入力により、1行ずつ上にスクロールしていく表示方法とした。実験は、6編の文書がランダムな順序で表示され、各々3回ずつ行われる乱塊法で行った。

(5) 測定：黙読開始から黙読終了までの時間を測定して、各回の黙読時間とその文書の全文字数から、1文字当たりの黙読時間を算出し、各文書3回の1文字当たりの黙読時間の平均を、平均黙読時間とした。

#### 3.2 黙読時間と RGV との相関

実験の結果得られた各文書の1文字当たりの平均黙読時間と RGV の関係を図2に示す。被験者間にばらつきが見られるものの、相関係数は0.65となり、正の相関が見られる。また、個人別の黙読時間と RGV との相関は、表2に示すように10名中7名に、0.9以上の高い相関が見られる。つまり、RGV の値が高くなるほど、1文字当たりにかかる平均黙読時間が長くなる傾向がある。

さて、各人のインタビュー結果から、文書の読み方は、様々であることがわかった。いつ、どこ、誰、といった要点をチェックする読み方、概要のみを把握する読み方等、読みの目的や態度が異なる場合や、飛ばし読みをする、文字を追って読む等、読むスタイルが異なる場合がある。このように、各人の読みの目的、スタイル、関心、知識などが異なるため、図2に示すように、同一の文書でも各個人の黙読時間にばらつきがでてしまったと考えられる。もし、実験条件として、読みの目的を定めたり、音読させて読みのスタイルを統一させれば、黙読時間のばらつきは減少し、

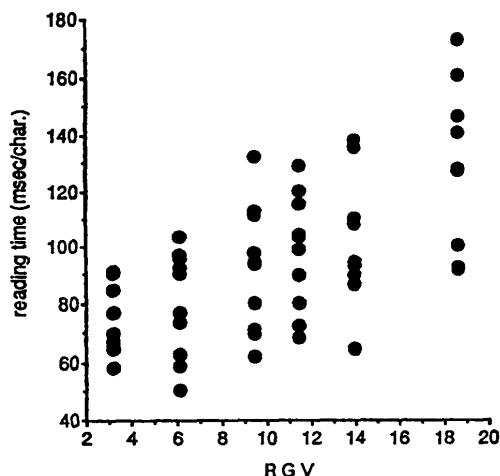


図 2 RGV と平均黙読時間の関係  
Fig. 2 Relationship between RGV and reading time.

表 2 各被験者の黙読時間と RGV との相関係数  
Table 2 The correlation coefficient between RGV and reading time for each subject.

| 被 験 者   | a     | b     | c     | d     | e     |
|---------|-------|-------|-------|-------|-------|
| 相 関 係 数 | 0.989 | 0.827 | 0.938 | 0.917 | 0.959 |
| 被 験 者   | f     | g     | h     | i     | j     |
| 相 関 係 数 | 0.603 | 0.945 | 0.971 | 0.709 | 0.969 |

RGV と黙読時間とのより高い相関が得られると考えられる。しかし、本実験では、黙読してくださいと指示する以外は実験条件を限定せずに、被験者に自由に黙読してもらった。このような条件でも、個人別の相関は極めて高く、さらに全体としても、相関関数が 0.65 と比較的高い相関が得られている。

このことから、読み手が誰であるかということと文書の RGV 値がわかれば、少ない誤差で黙読時間を予測することができると考えられる。さらに、読み手を特定しない場合でも、いくつかの文書の相対的な黙読時間の速さを、各文書の RGV 値から予測することができる。

したがって、RGV 値から予測した黙読時間を基準として、ディスプレイへのスクロール表示速度を設定することにより、読み手の黙読速度に適したスクロール表示方式を実現できると考えられる。読み手を特定しない場合に、読み手の黙読速度に適合していることを実証できれば、読み手を特定した場合には、さらに適合性が向上すると考えられる。そこで、次章では、読み手を特定しないスクロール表示について実験し考

察する。

#### 4. RGV を基準としたスクロール表示速度

ディスプレイに文書を表示する際の表示速度を、その文書の RGV 値に応じて設定する方式 (RGV 方式と呼ぶ) を提案し、従来の文書に無関係に速度一定の方式 (一定方式と呼ぶ) との比較実験を行い、読み手の黙読時間との適合性を検証する。

##### 4.1 RGV 方式と一定方式の比較実験

前章で文書の RGV 値とその黙読時間に高い相関があることを確認した。RGV 方式は、この相関関係を利用し、RGV 値からスクロール表示速度を求める関数をあらかじめ設定しておき、ある文書を表示する際に、まずその文書の RGV 値を(1)式で算出して、設定した関数によりスクロール表示速度を決定する方式である。

従来のコンピュータシステムでは、ディスプレイに表示するデータや文書は、可能な限り迅速に表示することが、暗黙の了解になっているため、モデムの通信速度やディスプレイの処理速度等の物理特性に依存して、表示速度が決定されていることが多かった。これに対し、RGV 方式で設定するスクロール速度は、読み手の黙読時間に依存して決めるため、従来方式に比べてより読み手の黙読時間に適合すると想定できる。

そこで、これを検証するため、文書と無関係にスクロール表示速度が一定の一定方式と RGV 方式を比較することとした。読み手の黙読時間との適合性は、目の動きから停留時間や逆行回数などを測定して検証することもできる。しかし、本実験では、各々の RGV の文書について、スクロール表示速度がどの程度適合しているかを定量的に検証することが必要である。そのため、設定された表示時間と、読み手が最適とする黙読時間を測定し、これらの時間差 (速度差) から適合性を考察することとした。

実験では、両方式ともに文書をディスプレイ画面から 1 行ずつ上にスクロール表示して、被験者に黙読してもらった。ただし、3 章の実験とは異なり自動的にスクロールする。被験者には、文書を黙読しながら、調節キーにより自分が最適とする表示速度に調節してもらった。なお、初期表示速度の設定は各々の方式で異なる。

RGV 方式の初期表示速度を、3 章の黙読実験の結果である図 2 から設定した。RGV 値と黙読時間の間には正の相関がある。さらに、RGV 値が低い文書に

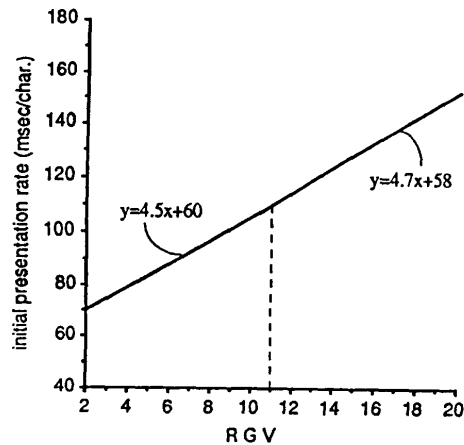


図3 RGV に応じて表示速度を設定する関数  
Fig. 3 The function of presentation rate based on RGV.

比べ、高い文書のほうが、黙読時間の変化の割合が若干大きくなる傾向が見られる。そこで、簡単のため図3に示す  $RGV=11$  を分岐点とする2つの1次式を用いることとした。この関数によると、RGV 値が2~11では、表示速度は RGV 1 レベル当たり 4.5 ミリ秒/文字変化し、RGV 値が 11~16 では、表示速度は RGV 1 レベル当たり 4.7 ミリ秒/文字変化する。一方、一定方式の初期速度は、図2の平均黙読時間（速度）である 110 ミリ秒/文字に設定した。

この2つの方式を比較する実験方法を以下に示す。

(1) 被験者：コンピュータのディスプレイを見ながらキーボードを使用することに習熟している成人6名であり、ランダムに3名ずつ、A、B 2つのグループに分けた。

(2) 文書：実験は、3章の黙読実験で用いた文書とは異なる新書4編、新聞2編、雑誌1編、教科書3編、計10編であり、RGV 値は、2.7~15.8 と幅広く分布している。文の長さは、1,075~1,570 字である。なお、10編の文書は、全被験者とも初めて読む文書である。

(3) 実験装置：3章と同様とした。

(4) 実験課題：被験者には、「調節キーを使って自分の読みやすい速度に調節してください」と指示した。被験者のAグループは、一定方式により表示された10編の文書を最初に黙読し、数日後 RGV 方式により10編の文書を黙読した。Bグループは逆に、RGV 方式を最初に黙読し、数日後一定方式により黙読した。10編の文書はランダムな順序で表示した。

(5) 測定：初期速度と速度調節後の最終速度との

表3 RGV に応じて表示速度を設定する方式 (RGV) と速度一定方式 (一定) との比較

Table 3 The statistics of reading times that were measured using the two types of text presentation controlling systems.

|       |      | (msec/char.) |    |     |     |    |    |
|-------|------|--------------|----|-----|-----|----|----|
| 調節量   | グループ | A            |    |     | B   |    |    |
|       | 被験者  | a            | b  | c   | d   | e  | f  |
| 標準偏差  | 一定   | 20           | 22 | 13  | 15  | 15 | 12 |
|       | RGV  | 12           | 12 | 8.5 | 8.3 | 15 | 13 |
| 絶対値平均 | 一定   | 16           | 21 | 13  | 21  | 16 | 16 |
|       | RGV  | 8.3          | 22 | 7.1 | 7.8 | 12 | 23 |

差（速度調節量）を測定した。

#### 4.2 RGV 方式の評価結果

各被験者の全文書に渡る速度調節量の絶対値平均と標準偏差を表3に示す。絶対値平均とは、全文書に対する速度調節量の絶対値を平均したものである。したがって、この値が小さくなるほど、初期速度が最適速度に近いといえる。標準偏差とは、各文書の速度調節量のばらつきである。このばらつきが小さいほど、どの RGV 値でも速度調節量が一樣であることを示しており、つまり読み手の最適とした速度と初期速度の差が、広い範囲の RGV 値の文書に渡って一樣であることを示している。

絶対値平均が小さく、かつ標準偏差も小さければ、本実験で与えた初期速度の値が、被験者の黙読速度に適合していたといえる。しかし、絶対値平均が大きい場合でも標準偏差が小さければ、RGV に対する表示速度の変化の傾きが、被験者の各文書に対する黙読速度の変化に適合していると考えられる。すなわち、初期速度の関数を上か下に若干シフトすることにより、被験者の実際の黙読速度に適合させることが可能であることを示している。

なお、最初に10編の文書を読んだ被験者が、2回目にも同一の文書を読むため、たとえ数日後であったとしてもその影響は残る。1回目と2回目の全被験者の絶対値平均の平均は、どちらも約15 ミリ秒/文字と同一であり影響は見られないが、標準偏差は、2回目のほうが1回目よりも平均して約3 ミリ秒/文字小さい。このため、標準偏差の値を用いて2つの方式を比較評価する際には、1回目と2回目による影響を考慮する必要がある。

さて、全被験者の絶対値平均の各方式での平均を比較すると、RGV 方式のほうが一定方式よりも約4ミ

り秒/文字 (約 1.6 秒/10 行) 小さいことがわかる。標準偏差は、Aグループで RGV 方式が一定方式よりも約 8 ミリ秒/文字小さく、Bグループで RGV 方式が一定方式よりも約 2 ミリ秒/文字小さい。Aグループの場合一定方式を初めに行ったので、RGV 方式と一定方式の差約 8 ミリ秒/文字には、1 回目と 2 回目による差 3 ミリ秒/文字の影響が含まれている。したがって、約 5 ミリ秒/文字 (約 2 秒/10 行) RGV 方式が一定方式よりも小さいと推定される。一方、Bグループの場合 RGV 方式を初めに行ったので、RGV 方式と一定方式の差約 2 ミリ秒/文字には、1 回目と 2 回目による差 3 ミリ秒/文字が負の影響を及ぼしている。したがって、Aグループの場合と同様、RGV 方式が一定方式よりも約 5 ミリ秒/文字標準偏差が小さいと推定できる。

以上のことから、被験者ごとに読む目的、スタイルなどが異なるため、個人間のばらつきは大きい、全体的に RGV 方式が一定方式よりも、被験者の黙読速度により適合した方式であるといえる。さらに、RGV 方式の初期速度設定関数を各被験者に適応させておけば、より適合度の高いスクロール表示速度を設定できると考えられる。

## 5. おわりに

本論文では、日本文の字面要因に対して因子分析と重回帰分析を行い、学年レベルを基準とした日本文の可読性の測度 RGV を、ひらがな、カタカナ、句点の頻度の 3 変数の線形式で定義した。また、コンピュータディスプレイ上で文書の黙読実験を行い、黙読時間と文書の RGV が正の高い相関を有することを確認した。この性質を利用し、RGV の一応用として、文書の RGV 値を基準としてスクロール表示速度を設定する RGV 方式を提案し、速度一定の方式との比較評価を行った。その結果、RGV 方式がより読み手の黙読時間に適合していることを示し、RGV の有用性を確認した。

本論文では、新たな日本文の可読性の測度を提案し、その新しい応用の可能性を考察した。しかし、読みの実験は、被験者に依存することが多いため、今後は、コンピュータの初心者など、被験者の数や層を増やして実験するとともに、読みの目的やスタイルなどの人に依存する要因と RGV 値との関係を検討する予定である。さらに、文書の読み書きの基準として、新たな応用方法を見つけることも大きな課題である。

**謝辞** 日ごろご指導いただく、ヒューマンインタフェース方式研究部遠藤隆也部長、徳永幸生グループリーダー、グループの皆様に感謝いたします。また、貴重など意見をいただきました青山学院大学理工学部矢頭俊介先生に感謝いたします。

## 参 考 文 献

- 1) Klare, G. R.: Assessing Readability, *Reading Res. Quart.*, Vol. 10, No. 1, pp. 62-102 (1974-75).
- 2) Flesch, R.: A New Readability Yardstick, *Journal of Applied Psychology*, Vol. 32, No. 3, pp. 221-233 (1948).
- 3) Dale, E. and Chall, J. S.: A Formula for Predicting Readability, *Educational Research Bulletin*, pp. 11-20 (1948).
- 4) Kincaid, J. P. et al.: Derivation of Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel, CNTT Research Branch Report 8-75, Naval Air Station, Memphis, Millington, TN: Chief of Naval Technical Training (1975).
- 5) Coke, E. U. and Koether, M. E.: A Study of the Match between the Stylistic Difficulty of Technical Documents and the Reading Skills of Technical Personnel, *Bell Syst. Tech. J.*, Vol. 62, No. 6, pp. 1849-1864 (1983).
- 6) 阪本一郎: 文章の語彙比重の査定法, *読書科学*, Vol. 6, No. 1-2, pp. 37-44 (1961).
- 7) 阪本一郎: 国語教科書の文の長さとその測定法, *読書科学*, Vol. 7, No. 2, pp. 17-24 (1962).
- 8) 阪本一郎: 文の長さの比重の査定法, *読書科学*, Vol. 8, No. 1, pp. 1-6 (1963).
- 9) 阪本一郎: 読みやすさの基準の一試案, *読書科学*, Vol. 14, No. 1-2, pp. 1-6 (1967).
- 10) 矢頭俊介: 邦文の可読性について, *青山経営論文集*, Vol. 11, No. 4, pp. 52-68 (1977).
- 11) 建石由佳, 小野芳彦, 山田尚勇: 日本文の読みやすさの評価式, *情報処理学会文書処理とヒューマンインタフェース研究会資料*, Vol. 18, No. 4, pp. 1-8 (1988).
- 12) 奥野忠一(編): 応用統計ハンドブック, 養賢堂 (1986).
- 13) 浅野陽子, 小川克彦: 読みやすい文書表示速度の検討, *人間工学*, Vol. 26, 特別号, pp. 234-235 (1990).
- 14) Mills, C. B. and Weldon, L. J.: Reading Text from Computer Screens, *ACM Comput. Surv.*, Vol. 19, No. 4, pp. 329-358 (1987).
- 15) 浅野陽子, 小川克彦: 日本文の読みやすさの一評価指標, 第 5 回ヒューマン・インタフェース・シンポジウム, pp. 113-116 (1989).

- 16) 浅野陽子, 小川克彦: 日本文の読みやすさに適応した文書表示速度の検討, 平2 信学会総全大, A-137 (1990).

(平成3年6月28日受付)

(平成3年10月3日採録)

## 付 録

因子分析の対象とする文書は, 次の20編である.  
一 小学校3年国語教科書 上・下 (光村図書, 昭和60年文部省検定済) より

「ありの行列」「イルカの会話」

「ニホンザルの仲間」「おにの話」「記号」

一 小学校6年国語教科書 上・下 (光村図書, 昭和60年文部省検定済) より

「自然を守る」「外来語と日本語」

「またとない天敵」「貝塚が教えるなぞ」

「フェアプレー」

一 中学校3年国語教科書 (光村図書, 昭和61年文部省検定済) より

「動物の文化的行動」「広く学ぶ心」

「新しい文体の誕生」「民族と文化」

「人間の叡知を」

一 高等学校3年国語教科書 (尚学図書, 昭和55年文部省検定済) より

「国語の課題」「旅の思想」「ナマズ考」

「雪国の春」「地球の安全」

以下に, 実際の文書の例を挙げる.

### 小学3年の文書の一例「ありの行列」の一部

夏になると, 庭のすみなどで, ありの行列をよく見かけます. その行列は, ありのすからえさのある所まで, ずっとつづいています. ありは, ものがよく見えません. それなのに, なぜ, ありの行列ができるのでしょうか.

アメリカに, ウィルソンという学者がいます. この人は, 次のようにして, ありのようすをかんさつしました.

はじめに, ありのすから少しはなれた所に, 一つまみのさとうをおきました. しばらくすると, 一びきのありが, そのさとうを見つけました. これは, えさをさがすために外に出ていたはたらきありです. ありは, やがてすに帰っていきました. すると, すの中から, たくさんのはたらきありが, 次々と出てきました. そして, 列を作ってさとうの所まで行きました. ふしぎなことに, その行列は, はじめのありがすに帰

るときに通った道すじから外れていないのです.

### 中学3年の文書の一例「動物の文化的行動」の一部

「人間は文化をもった動物である.」といわれている. 確かに, 文化は人類社会の特徴であり, 文化なくして人類社会の成立は考えられない. しかし, 動物には全く文化現象がないといいきれののだろうか. この問題に明確な答えを出したのは, 日本の霊長類学者だった.

嵐山や幸島のニホンザルは, 小鳥の卵を取って食べる. ところが, 高崎山や小豆島の群れのサルは食べない. 与えてみると, 全く無視するか少し触ってみるだけである. こういう例は, ほかに幾つも見られる. 幸島のサルはマサキの実を食べるが, 屋久島の西北海岸の崖にすむサルは食べないとか, 志賀高原や白山の積雪地のサルは冬には木の皮をよく食べるが, 南国のサルは食べないといったことが見られる. こういう現象を, 群れによって食文化が違っていると表現してはいけないのだろうか.

### 高校3年の文書の一例「国語の課題」の一部

標準語とは, 一口に標準語といっても, 二つの意義がある. 京都の旅館で, 女中が「お客様へは京都弁ではなく, 標準語を使うようにいたしております.」と言うときの標準語は, 東京弁即標準語であって, いわば, 中央方言・共通語の意味を出ない. 東京語は現に, 全日本の共通語となりつつある中央方言である.

また, 放送局のニュースの報道には, なるだけ標準語でと努力している. この場合には, もはや単なる東京弁という意義ではない. ここにいう標準語とは, 特に現代日本語の典型的な, 代表的な, 標準的な, できれば模範的な, 発音や用語を期待してそういつている.

今問題にする標準語は, 後の意味の標準語である. では日本の国語をよくするために, 手本とすべき標準語をどう定めるべきか, ということを考えてみよう.



**浅野 陽子**

1965年生。1988年慶応義塾大学  
理工学部管理工学科卒業。同年より  
NTT ヒューマンインタフェース研  
究所に勤務。文書の読みやすさの研  
究に従事。電子情報通信学会，日本  
人間工学会各会員。

**小川 克彦 (正会員)**

1954年生。1976年慶応義塾大学  
工学部管理工学科卒業。1978年同大  
学院工学研究科修士課程修了。同年  
日本電信電話公社入社。現在，NTT  
ヒューマンインタフェース研究所主  
幹研究員。工学博士。画像情報処理システムの研究開  
発を経て，現在，ヒューマン・コンピュータ・インタ  
ラクション，認知工学の研究に従事。電子情報通信学  
会，日本人間工学会，ヒューマンファクターズソサエ  
ティ，IEEE 各会員。