

文書の彩色による可読性向上の評価

内田 友幸、田中 英彦 (東京大学大学院工学系研究科)
{tomo, tanaka}@mtl.t.u-tokyo.ac.jp

1. はじめに

近年、電子化された自然言語文書が大量に流通するようになってきたが、これらの自然言語文書に対する UI は従来の紙メディアと同様の白黒の表現が主流である。しかし、カラーディスプレイなどのデバイスのカラー化に伴い、自然言語文書に対しても色を利用した有効な UI が望まれてきている。そこで我々は文書を彩色する事で読みやすくする検討を重ねている。本稿では、彩色の影響を心理実験を用いて調査を行なったので、その実験の内容と考察について報告する。

2. 心理実験概要

心理実験は 2 種類行なった。一つはアクセス速度実験で「文書中の要素に対するアクセスの高速化」に着目し、彩色による速度の差について評価した。もう一つは読解実験で、「文書の概要の把握」と「読む速度」の点に着目し、色情報の付加による文書の読解速度と理解度を評価した。

アクセス速度実験はまず、新聞記事中の単語一つを赤字または太字で強調したものを用意し、ディスプレイ上に表示する。被験者には強調された単語の品詞による分類を行なってもらい、それにかかった時間を測定する。これにより、単語の強調の仕方によるアクセス速度の違いを測定できる。

読解心理実験は文書を提示し、読み終わったらキーを押してもらい、内容に関する質問に答えるという手順で行なった。白黒の文書と彩色文書それぞれについて読むのにかかった時間、質問の正答率、答えるのにかかった時間を計測し、白黒の時と彩色した時との差異を評価する。

3. アクセス速度心理実験

3.1 文書提示方法

被験者にはビットマップで文書を提示する。表示には 17 インチ CRT を利用し、被験者の目の位置から 1m 離れた位置に暗所にて表示させた。

各文書毎を約 1200 × 500 ピクセルのビットマップで表示し、文字は明朝体で記述してあり、太字は約 2 割ほど大きくしたゴシック体で指定している。この際、表示される文字の 1 辺は 6mm であるので、視角にして 0.34 ° に相当する。

読解実験でも同様に文書を提示するが、被験者の目の位置を 80cm と若干近付けた。

なお、この CRT の輝度や色座標などの特性をミノルタの CL-100 によって計測した。この結果

を読解実験の条件と共に表 1 に示す。

表 1. CRT 表示の輝度と色座標

アクセス実験	輝度 (lx)	色度 (CIE 1931 XYZ 表色系)
白 (文字の背景)	38.5	(0.292, 0.323)
赤 (赤字)	6.9	(0.623, 0.341)
読解実験		
白 (文字の背景)	42.8	(0.273, 0.295)
灰 (記事の周囲)	22.7	(0.277, 0.289)
紺 (どうだ、どうした)	1.7	(0.145, 0.064)
青緑 (片仮名 Alphabet)	5.9	(0.205, 0.279)
紫 (数字)	4.0	(0.183, 0.084)
緑 (固有名詞)	8.4	(0.299, 0.616)
灰 (上記以外の平仮名)	7.7	(0.279, 0.297)
黒 (上記以外の漢字)	0.0	-

3.2 被験者の操作

被験者は、文書中に一箇所赤字か太字で強調された 2 字熟語の単語を固有名詞かサ変動詞性名詞のどちらかを判別し、2 つのキーのどちらかを押すことで分類を行なう。

もう一つの操作は、強調された単語を発見するだけというもので、なるべく早く探し、発見した時点で指定のキーを押してもらった。

まずは 40 記事を品詞毎に分類し、次に同じ 40 記事中の強調された単語を探索してもらった。

3.3 アクセス速度実験結果

被験者の数は 6 名。統計的解析の結果を表 2 に示す。

表 2. 単語の分類時間

	赤字	太字
平均	941	1.24x10 ³
標準誤差 (95%)	21.2	38.1

この t-検定の結果は 8.08 であり、この 2 群は分離しているといえる。これより赤字の方が単語の探索、認識にかかる時間は少なくなると言える。

分類に要した全時間から単に単語の探索をして押すだけに要した時間を引くと単語の認識所要時間に相当すると考えられる。

この認識所要時間は平均で赤字が 604ms、太字で 462ms、t-検定の結果が 2.87 なのでこの 2 群も分離している。この結果から赤字の方が太字よりも認識の所要時間が大きいことがいえる。

4. 読解心理実験内容

4.1 文書提示方法

被験者には一回の測定で文書、質問文それぞれ

一つずつを順次提示する。文書は小説、紀行文の意味的にまとまりのある部分からそれぞれ約 120 文字を切りだし、40 文字毎に 3 行にして 954 × 153 ピクセルの一枚のビットマップにした。

質問文についても同様に約 20 文字で構成したビットマップとした。質問文は文書の内容に関わることを簡潔に述べた一文によって構成され、その半数については虚偽の記述となるように設定した。

彩色文書の彩色ルールは文字種 (片仮名とアルファベット、数字、平仮名、漢字) によるものと固有名詞、文が述べている内容の「どうだ」「どうした」にあたる述語の 6 種類に分類した上で、それぞれに色を決めて、文字色として彩色した。

4.2 被験者の操作

被験者にはまず文書を提示し、読み終わったらキーを押してもらおう。キーが押されたら今度は質問文を提示し、その正誤を判断して相当するキーを押してもらおう。このそれぞれの、表示からキーの押下までの時間と正誤に関する正答率を計測する。

利用した文書は全部で 36 文で 6 文はトレーニング用とし、30 文を 15 文ずつに分け 1 群、2 群とする。被験者には 1 群の白黒文書と 2 群の彩色文書の組合せか、1 群の彩色文書と 2 群の白黒文書の組合せのどちらかを順次提示する。

4.3 読解実験結果

被験者の数は 12 名で、各群毎に 15 個の素データを得た。文章を読み終るまでの読解所要時間を表 3 に示す。

表 3. 読解所要時間

諸元	1 群彩色	1 群白黒	2 群彩色	2 群白黒
平均 (ms)	1.213x10 ⁴	1.016x10 ⁴	1.033x10 ⁴	1.158x10 ⁴
95%標準誤差	530	398	408	477

1 群では白黒の方が速く、2 群では彩色の方が速いという結果になった。これは 1 群白黒と 2 群彩色を行なったグループに読解速度が速い人が何人かいた影響が大きいと考えられる。グループ間の 15% のこれらの影響を加味して判断すれば彩色と白黒の時の読解所要時間に有意な差があるとはいえない。

表 4. 回答所要時間

諸元	1 群彩色	1 群白黒	2 群彩色	2 群白黒
平均 (ms)	2288	2196	2352	2743
95%標準誤差	126	122	108	162

続いて回答にかかった時間を表 4 にまとめた。これも 1 群では白黒の方が速く、2 群では彩色の方が速いという結果になった。これもグループ間の 11% の影響を加味して判断すれば有意な差があるとはいえない。

次に誤答数をグラフに表したものを図 1 に示す。

図 1 誤答数

誤答数を比べると合計で白黒が 25 に対し、彩色が 14 と 44% も低くなっている。また、彩色の方が誤答数が少ない人が 6 名、白黒の方が誤答数が少ない人が 2 名で人的的にも彩色の方が誤答が少ない方向にある。

5. 考察

実験の結果から文書中の読ませたい単語を赤色に彩色することは太字で指定するより所要時間の短縮につながるということが言える。しかし、この短縮は主に単語探索時間の短縮によってもたらされているものであり、見つかった単語を認識するのにかかる時間は逆に赤字の方が大きくなってしまっている。これは太字よりも赤字の方が読みにくいということを定量的に示していると言えるだろう。この読みにくい原因には赤が黒よりも輝度が高いことによるコントラストの低下なども考えられるが色情報が読解過程を妨害していることも十分に考えられる。被験者の感想の中に赤いとチラチラするというものがあるが、これがこの読解過程の妨害に該当しているとも考えられる。

読解実験の結果から文書を彩色してやることにより設問に対する正答率が上がるが、読む時間は白黒文書と変わらないという事がいえる。設問は単に色の分布を見るだけでは回答できず、文章の意味を理解している必要がある。そのため、正答率が上がるのは彩色文書の方が文書の内容をより正確に理解し、覚えているということが推察される。被験者の弁を借りれば「イメージが頭に浮かびやすく残りやすい」という感覚のようだ。

文書の彩色には弊害も懸念されるが、多彩な可能性があり、うまく利用できればかなり有効であると考えられる。今後、自動彩色システムの構築などを行ない、さらに検討を重ねていきたい。

文 献

内田 友幸 田中 英彦: “文書の概要把握過程における彩色の有効性の評価” 情処第 52 回全国大会, 5D-5(March 1996).