

HyperTextのReadability評価

杉原 敏昭 嶋田 敦夫

(株) リコー 中央研究所 知能工学研究センター

分かり易く、使いやすい計算機システムを目指して、いわゆる”計算機のメディア化”の研究が広く行なわれている、我々はそのメディアとしての計算機を扱う人間の側のファクターを定量的に解析して、より洗練された計算機の操作環境を実現することを目標に研究を進めている。

今回の報告では、代表的なハイパーテキストシステムであるHyperCard™上に構成された文書と紙の文書の読解性の比較を心理学実験を用いて検証した。読解性は、文書の読み取り速度及び内容に関する記憶テストで測定された。その結果、読み取り速度の点ではハイパーテキストは紙よりも劣る、しかし、テキストの読後に内容の記憶テストを行なうとハイパーテキストは紙よりも優位な結果を示した。すなわち、再認テストでは両方の文書で差異が認められないが、再生テストではハイパーテキストが良好な成績を示した。

A Psychological Evaluation for Readability of HyperText

Toshiaki SUGIHARA Atsuo SHIMADA

Artificial Intelligence Research Center, RICOH R&D

16-1 Shin'ei-chou, Kouhoku-ku, Yokohama 223, Japan

ABSTRACT: This paper describes how HyperText system affects readers' memory processes and resulting readability. Using psychological experiment, we compared and evaluated a readers' comprehension task, under two different condition; reading plain paper text document and reading HyperText document on HyperCard™ system. We measured two evaluation factors, one was reading time for entire document, another was result score of examination for its contents. As a result of that, we got following two evaluation;

- (1) Reading time analysis shows, HyperText made much longer reading time than plain paper text, that was already reported by prior researchers in this field, we also verified.
- (2) In score analysis, we focused its on both recall and recognition memory processes. From both point of view, in recall case, it was significant difference between HyperText and plain paper text. Above those results suggests that HyperText is a useful tool for subjects to encode information more easily.

はじめに

情報化社会が進むにつれて、ますます計算機と人間のコミュニケーションは頻度と範囲を増している。その結果、ヒューマンインターフェースは重要な研究課題であると考えられている。現在、われわれは情報伝達能力に優れたハイパーテキストを用いたCAI (Computer Aided Instruction) 研究を進めているが、そこでも、人間の計算機利用のしかたに関する問題が提出される。そのひとつに計算機による情報の提示がある。これに対し、われわれは、人間が計算機画面に表示されているテキストを読むという行動に着目して、計算機で文章を提示する際の問題を実証的に検討し、計算機にとって有効な提示方法を探ることが必要であると考えている。

そこで、われわれは題材として、本研究では文字や絵を用いることのできるハイパーテキストシステム上に構築されたテキストを用い、これと従来の紙の上に書かれたテキストの読解性 (readability) を心理学実験を通じて比較することによって上記の点について検討した。実験の結果、計算機に有利ないくつかの特性が示唆されたので報告する。

読解性

メディアとしての計算機が人間の文章を読む過程に与える影響は、一般に可読性 (legibility) と読解性 (readability) から調べられる。可読性とは文字や単語自体がどの程度識別し易いかを示すものである。通常、文字や単語にノイズを含めて提示し判読可能かを調べることによって測定される。一方、読解性は、文字単位ではなく文章全体の内容がどの程度理解し易いかを示すもので[1]、文章の読み取り速度や被験者がどの程度内容を理

解したかを調べることによって測定される。

計算機におけるテキストの読解性は、紙のテキストと比較することによって測定されている[2][3][4][5]。その方法は概ね、内容が同一のテキストを紙とCRT画面に呈示し、文章の読み取り速度や理解作業のパフォーマンスを比較することである。その結果、文章を校正する作業ではCRT画面のテキストと紙のテキストとはパフォーマンスが同等、もしくは紙の方が速く正確に行なえることが示されている[2][3]。またテキストを読む速度に関しては紙面の方が速いことも示されている[4]。このような結果の要因として[6]は被験者のCRT画面への注視角度、視野、画面のちらつきなどを考えたが、現在のところ、その要因は明らかではない。しかしながら文章校正やテキストの読み取り速度と言った比較的単純な作業の指標においてはCRT画面の文章が劣ることはわかっている。

[5]は技術マニュアルを呈示しスイッチ設定作業という問題解決課題を行なわせ紙とCRT画面の読解性を検討した。[2][3][4]と異なりこの実験はかなり複雑な課題を使用しているが、同様に設定作業に要した時間はCRT画面のマニュアルの方がより長くなった。

このように、課題の複雑さにかかわらずCRT画面が紙に比べ読解性で劣ることが示されている。被験者はCRT画面の方でより長い時間をかけてテキストを読むが、理解度としては紙のテキストと同等もしくは劣るということになる。このうち理解度については、実験が理解度のどの側面を測定したのが明確でないことから、CRT画面が理解度を低減させるという結果に対してさまざまな議論がなされている。例えば、[1]は理解度を測定するテストが単純すぎて紙面表示と画面表示の差の要因が同定できないのであろうとしている。

このような背景から、読解性の測度が理解のどの側面を測定するものなのかを規定した上での考察が必要となろう。本実験ではわれわれは理解過程の一側面をテキストの内容を記憶する過程であ

ると捉え直した、それゆえ読解性は、記憶された情報が後にどの程度想起されるかということで測定できると考えられる、そうすることで、「なぜ計算機というメディアが人間の読解性に影響するのか」という問に対し、記憶の側面から答えられる筈である。また、もし再生・再認の測度が計算機と紙の比較において異なった結果を示したならば、今後、計算機が強調しなければならない側面も明らかになるはずである。ところで、人間が情報を想起する形態には、再認 (recognition) と再生 (recall) がある。再認とは記憶された情報があとで呈示された情報と同じものかを判断することで、ターゲットの単語そのものを想起することはできないがそれをリストの中から選ぶことができることを言う。一方、再生は記憶された情報そのものをあとで想起できることを言う。例えば、思い出したい旧友の名前はわからないがその名前を聞けば正しいか否かがわかることは再認にあたり、思い出したい旧友の名前そのものを想起することが再生にあたる。このような観点から読解性を捉え直し、もし再認・再生の測度が計算機と紙との比較において異なった結果を示したならば、今後、計算機が強調しなければならない側面が明らかになると思われる。

以上から、われわれは次のような目的を設定し、これらを実験を通じて検討していく。

1. 従来の研究と同様にハイパーテキストを用いても紙のテキストより読解性が劣るかを確認する。また劣るとすればどのような側面で劣るのかを同定する。
2. 読解性を再認・再生という枠組みで捉えると紙のテキストとハイパーテキストとで成績に差が現われるかを確認する
3. 差が現われるとすれば、その枠組みにおいて言えるハイパーテキストの優位性はなにかを検討する。

方法

紙のテキストと計算機上のハイパーテキストの読解性を記憶過程という観点から比較するため、紙のテキストを読む条件と計算機上のハイパーテキストを読む条件が設けられた。各条件毎に被験者はテキストを十分読まされ、テキストを読む時間、テキストの内容の再認数、再生数が測定された。

材料

テキスト 米国アップル社のMacintosh上で稼働するハイパーテキストシステムHyperCard™上に構築されたハイパーテキストの入門スタックを計算機条件の刺激として使用した。またその画面イメージをA4版にプリントアウト (43ページ) したものを紙条件のテキストとして使用した。図1にこの実験で用いたスタックでの基本的なカードの構成を示す。この入門スタックでは、リンクは基本的に連続したカード間を結ぶシークエンシャルな構造である、一部には記述のより詳細な説明等の枝葉分れしたリンクを持つが、これらでも、その記述を読んだ後には必ず元のカードに戻る様に構成した、即ち、読み方の順路は一意である。両テキストから与えられる情報を条件間で同一にするため、紙条件のテキストもカード内のボタンをクリックしたときの画面情報を含んでいる。

再認・再生テスト 再認テストは正誤判断の形式をとっており、テキストに答えのあるものとテキストを読んでも答えられないもの (ダミー) が用意されている。図2に再認テストの課題の一部を示す、質問数はダミーを除くと27である。

再生テストはテキストに記述されているハイパーテキストシステムの操作方法を再生させるものである。計算機上で実際にハイパーカードを操作させることで測定した。図3に再生課題の一例を示す、これら再生課題は全部で10問用意した。

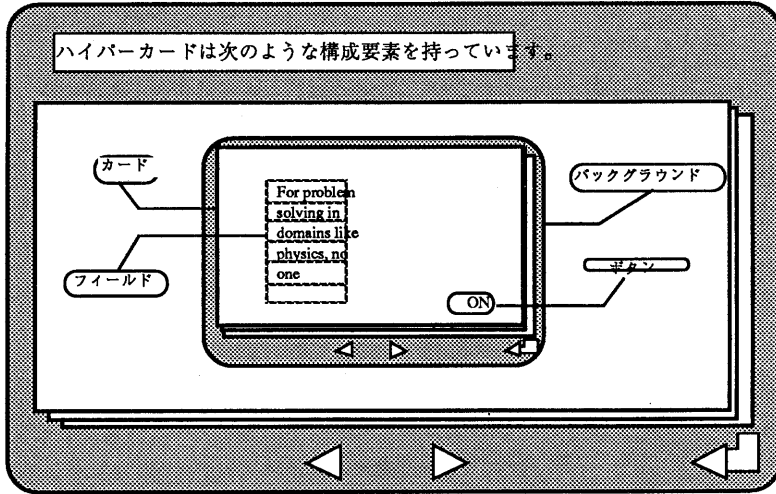


図1：実験に用いたカードの基本構成

次の文章はハイパーカードについて記述したものです。内容が正しいものをすべて選び出して下さい。

- 1.カードが束になったものをスタックという。
- 2.フィールドの形状や大きさは自由に設定することができる。
- 3.カードの大きさは自由に設定することができる。
- 4.ユーザーの使用レベルは全部で10段階ある。
- 5.ハイパーカードは一種のプログラミング言語である。
- 6.ユーザはブラウジングのときにメニュー

図2：再認テストの課題例

次の作業を実際にコンピュータを使って行ってください。

The diagram shows a stack of cards. The top card is numbered "1" and contains the following text:

ホームカードのバックグラウンドのスク립トを読みだしてください。

 The stack is shown with several other cards underneath, indicating a sequence of tasks.

図3：再生テストの課題例

被験者

(株)リコー・中央研究所の研究員24名(23-38歳)。全員、ハイパーテキストに対する初学者であるが、マウスやアイコンの操作などの計算機に関する知識を多少もっている。

手続き

教示 テキストを読んでハイパーテキストシステムについて理解を深めるよう教示した。通常の目的でテキストを読んでもらうため、被験者は実験後にテストが実施されることは告げられなかった。また、計算機条件では操作方法(被験者が操作できる機能はページ操作アイコンとハイパーテキストに埋め込められているボタンのみである)に関して最小限の解説をした。

手続き 口頭で教示を行なった後、実験者は被験者にテキストを読むよう指示した。開始の指示から被験者の読み終わりの合図までを読み有した時間として測定した。読後、再認テスト、再生テストの順に読解性の測定を実施した。再認テストでは、被験者は正誤判断に対する確信度の自己評定を求められた。再生テストでは、被験者が3分以上経過しても課題を達成できないときは、課題達成の見込をきき、見込のないと申告した場合

は、その課題を打ち切った。さらに被験者が正解に達しても操作方法を記憶していたかを聞いた。

結果

われわれが本実験で測定した従属変数は、テキストを読むのに有した時間及び再認テストの成績、再生テストの成績である。このうち、再認テストはダミーを除く正誤判断のうち正解であった質問の数をカウントした。なお、確信度の自己評定において自信がほとんどないと申告したものは被験者の判断に信頼性が欠けるのでカウントから除外した。また、再生テストでは達成した操作課題をカウントしたが、偶然による正解は除外した。

これらの測度に対して紙のテキストを読む群と計算機上のハイパーテキストを読む群でそれぞれ12名づつ、計24名のデータが測定された。ただし、被験者総数が24名と少数で母集団が正規分布に従う保証がないので、上記測度に対して条件間で分散に差がないことを検討した。その結果(F値)、及び、平均値と分散を表1に示す。

結果：記憶テスト(再認テストのみ) $t=0.721, df=22, p>0.40$

F(11,11)=1.959	ハイパーテキスト	紙のテキスト
平均値分散	22.083 14.410	21.250 7.355

結果：記憶テスト(再生テストのみ) $t=2.167, df=22, p<0.025$

F(11,11)=1.70	ハイパーテキスト	紙のテキスト
平均値分散	3.750 1.354	2.167 2.306

表1

結果：テキストを読む速度 $t=3.534, df=22, p<0.005$

F(11,11)=1.541	ハイパーテキスト	紙のテキスト
平均値分散	1292.58 73425.6	921.833 47642.5

表2

表1からすべての測度に対して条件間で母集団に差のないことが認められたので、各々に対して条件間の差の有無をT検定を実施して調べた。結果を表2に示す。T検定の結果、テキストを読むのに有した時間に有意差が認められた ($t=3.534, df=22, p<0.005$) (表2)。つまり、被験者は紙のテキストを読む方が計算機上のテキストを読むより速い傾向があるといえる。一方、再認・再生を個別に検定すると、再認テストでは統計的有意差が認められないが、再生テストにおいて若干の有意差が現われる ($t=2.167, df=22, p<0.025$)。本実験で用いたテキストは、再認数には紙と計算機で差がないが、操作方法を再生するという課題では計算機上に書かれたテキストの方が好成績を示したことになる。つまり、被験者はハイパーテキストの方がより多くの内容を再生できたと言える。

考察・結論

本実験において、われわれは読解性をテキストの読み取り速度と理解度から測定した。理解度については特に理解過程の一つの側面をテキストの内容を記憶する過程と捉え、再認テストと再生テストを実施し、検討した。その結果、従来の研究

と同様に被験者は紙に比べより長い時間をかけて計算機上のテキストを読む傾向にあった。

また、理解度の測定ではハイパーテキストを計算機画面に表示したものの方が再生テストにおいて良好な成績を示した。そこで、再認テストに差が現われず再生テストに差が現われたことについて以下に考察する。

新しい情報が記憶されるとき人間は常に注意を向けていないと短期記憶の活性は弱まっていき、情報の再生ができなくなってしまう。そこで、ある期間情報を保持するためリハーサルという短期記憶の活性維持の試みを行なっている。Craig, F.I. M & Lockhart, R.S. [7]によればリハーサルには2種類ある(図4)、ひとつは維持リハーサル (maintenance rehearsal) とよばれ単に短期記憶内に情報を保持しておくためのものである。もうひとつは精緻化リハーサル (elaborate rehearsal) で、記憶に留めるべき情報の意味や他の情報との関連を分析して保持するものである。さらに、Woodward, A.E., Bojork, R.A. & Jongeward, R.H. [8]は精緻化リハーサルは再生のしやすさを増大させ、維持リハーサルは再生のしやすさには影響を与えないが、再認のしやすさには強い効果をもっているとしている。

このことから計算機上にハイパーテキストを呈示した条件の被験者は、紙のテキストを呈示され

2種類のリハーサル

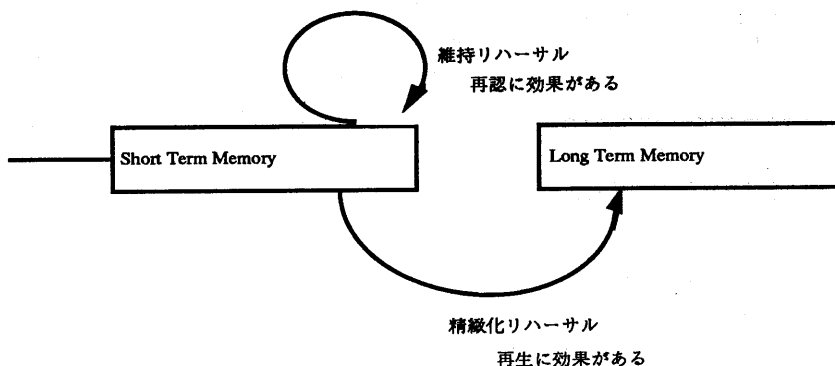


図4：リハーサルの概念図

た場合より精緻化リハーサルが行ないやすい環境にあったと推察される。言い換えれば、ハイパーテキストは情報の意味づけや既存知識との関連をリンクを用いて動的に行なえるため、従来のCRT上のプレーンなテキストや紙よりも符号化(encoding)が容易なメディアであるといえよう。

また、一般に、再生は再認に比べて記憶検索のための手掛かりが非常に少ない。なぜならば再認テストでは想起すべき情報という強力な手掛かりが直接与えられるからである。それゆえに被験者は再生の場合、多くの手掛かりを自発的に生成する必要があると言える。ハイパーテキストは再生に対する精緻化と手掛かり生成の要件を備えていると考えられるが、反対にハイパーテキストで文書を構成する場合には、ユーザに理解させたい内容が意味処理しやすい様に設計に留意する必要があるとも言える。

謝辞

本研究を行なう機会を与えて頂いた森後二所長ならびに実験に協力して頂いたリコー中央研究所の所員に深く感謝します。

参考文献

- [1] Mills,B.M. & Weldon,L.J. 1987. Reading Text from Computer Screens. ACM Computing Surveys,Vol. 19,No.4,pp.329-358
- [2] Wright,P. & Lickorish,A. 1983. Proof-reading Texts on Screen and Paper. Behavioral Information Technology,Vol.2,pp.227-235
- [3] Gould,J.D. & Grischkowsky,N. 1984. Doing the same work with hard copy and with CRT computer terminals. Human Factors, Vol.26,pp.323-337
- [4] Muter,P., Latremouille,S.A., Treurniet,W.C. & Be am,P. 1982. Extended reading of continuous text on television screens. Human Factors,Vol.24,pp.501-508
- [5] Weldon,L.J., Mills,C.B., Koved,L. & Shneiderman,B. 1985. The structure of information in online and paper technical manuals. In R.W.Swezey(Ed.) Proceedings of the Human Factors Society 29th Annual Meeting,Vol.2,pp.1110-1113
- [6] Gould,J.D., Alfaro,L., Barnes,V. Finn,R., Grischkowsky,N. & Minuto,A. 1987. Reading is slower from CRT displays than from papers: Attempts to isolate a single-variable explanation. Human Factors,Vol.29,pp.497-517
- [7] Craik,F.I.M. & Lockhart,R.S. 1972. Levels of Processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior,Vol.11,pp.671-684
- [8] Woodward,A.E., Bojork,R.A. & Jongeward,R.H. 1973. Recall and recognition as a function of primary rehearsal. Journal of Verbal Learning and Verbal Behavior,Vol.12,pp.608-617