

視線情報を用いたレトロスペクティブインタビュー法の開発とその効果

The development of Retrospective Interview method with Eye movements and effectiveness of this method

湯澤 秀人†

Hideto Yuzawa †,

戸崎 幹夫†

Mikio Tozaki †

1. はじめに

現在、技術の進歩に伴う多機能化やユーザーニーズの多様化に伴い、機器のインターフェースがより複雑となっており、その使いやすさが非常に重要な問題となっている[1], [2]. これに対して各メーカーや外部評価機関では、どこに、どのような問題が発生しているのかなどユーザービリティに関する調査を実施し、使いやすさ向上を図っている。特に近年のユーザービリティ調査では、ISO9241-11に代表されるようにタスク成功率、エラーの数や種類、主観評価など、人間側に重きをおいたパフォーマンス測定を重視する傾向となっている[3].

確かにこれら計測は非常に重要であるが、本質的なユーザービリティ向上のためには、これらに加えて、操作プロセスを示すユーザーの認知的側面を捉える必要があると考えられる。例えば、結果的にタスクを達成できたとしても、ユーザーの思い通りに操作していたのか、それとも機器主導であったかなどの潜在的な問題が挙げられ、これらを考慮してデザインすることによって初めてユーザーにとって使いやすい、「透明な」インターフェースを実現することができる。

この問題を解決する手段として、ユーザーの操作中に利用される発話思考法や操作後に利用される半構造化面接

などの従来手法が挙げられるが、操作中のユーザーの思考を干渉してしまうこと、ユーザーの想起の手がかりが少ないために、操作後の記憶の再生が困難であるなどいくつかの課題があった。

本論文では、視線情報を用いた新しいインタビュー手法を提案し、更に従来手法との効果比較について報告する。

2. インタビュー手法の提案

ここで提案する手法を Fig.1 に示す。まず評価対象となる機器を被験者に操作してもらい、操作中の被験者の視線情報を取得しておく。勿論、操作中に被験者に視線情報が提示されることはなく、被験者は通常の操作を行うのみとする。操作後、「操作中の視線情報と操作情報を合成したビデオ画像をダイナミックに再生しながら」、被験者へのインタビューを実施し、被験者に各操作において何を意図して、どのように操作しているかを発話させる。

すなわち、通常、ユーザーには意識されることのない視線軌跡を想起の手がかりとして用いることによって、操作中のユーザーの思考を干渉することなく、事後のインタビューを行う手法である。更に、これによって、被験者からの発話に有効な変化が見られるならば、本手法は従来手法と比較して効果的であると結論付けることができる。なお、計測可能な生体情報としては、視線以外にも脳波や心拍などいくつか考えられるが、インターフェース上の行動を直接的に表現している視線情報をまずは用いることにした。

以下では、この手法が実用上、適用可能かどうか、そして視線情報を用いない従来手法と比較して有効であるかどうか、有効であるならばどこに違いが表出するか検証した。

3. 検証実験

3.1 実験仕様

被験者は、男性6名、女性3名の計9名、年齢は30歳から45歳(平均35.1歳、SD=4.9)を対象とした。職業は、業務上でPCを利用するオフィスワーカーであり、PCの操作には十分慣れていた。

実験条件は、レトロスペクティブインタビュー時に被験者に提示するビデオ画像の視線付加条件(視線あり条件、視線なし条件)のみとし、各被験者はどちらか一方の実験条件

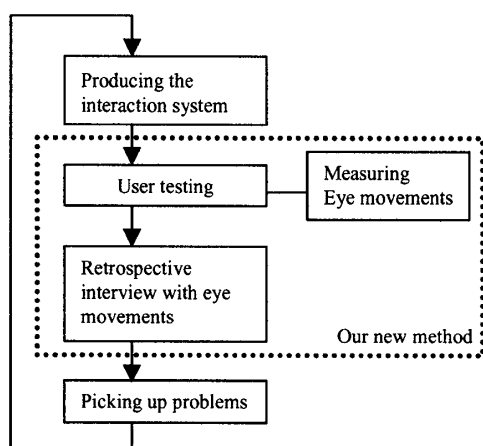


Fig.1 The retrospective interview method with eye movements

のみが与えられた。なお、視線計測は両条件とも実施した。この実験条件において、各条件の発話量、発話内容、実験に対する主観評価について比較した。

対象刺激は、開発途中のプリンタドライバとした。これによって、「開発途中のプリンタドライバの市場投入前テストのため、思ったことは存分に言って欲しい」と、被験者が話しやすい自然な状況を設定し、実験意図の察知問題を考慮した。なお、実験後のヒヤリングより、本来の研究の意図に気づいた被験者はいなかった。

3.2 実験手順

3.2.1 インTRODクシヨ

被験者へ実験手順に関する説明を行い、更に事前に収録した視線情報を付加した操作中のビデオ画像を再生して、視線情報に関する基礎知識を説明した。これは、視線情報の基礎知識を持たない被験者の場合、視線情報自体の理解に負荷がかかり有効な発話を得られない、あるいは視線情報の間違っ了解による発話を防ぐ、といったことを配慮するため実施した。但し、事前のこの説明自体が、その後のインタビューに影響を与えることも考えられるので、両条件において同質の説明を行った。

3.2.2 タスク実施

タスクは、実験手順や操作に慣れるためのアイドリングタスク(タスク1)を1つ、事前実験により難易度が異なったタスク(タスク2,タスク3)の2つ、計3つのタスクを実施した。タスクの概略としては、タスク1がB4サイズのWord文書をA3用紙に出力する、タスク2がA4サイズのPowerPoint文書をサンプルと同じようにレイアウトして出力する、タスク3がはがきサイズのWord文書を手差しトレイで出力する、といったものである。なお、タスクの実験順序は、タスク1のみ全被験者で固定し、タスク2,3はランダムに入れ替えて実施した。

3.2.3 レトросペクティブインタビュー実施

全てのタスクが終了した後、前述の各実験条件のビデオ画像を再生しながら、レトросペクティブインタビューを実施した。但し、インタビューの聞き方が、発話量、発話内容等に強く依存すると考えられるため、ここでは、あえてインタビューは聞き役に徹し、全被験者には「これから再生するビデオ画像を元にご自分の操作を逐次、説明していただき、不満や思ったことなど何でもおっしゃってください」と指示した。その際、必要があれば適宜、一時停止をしながらインタビューを行った。

3.2.4 主観評価

レトросペクティブインタビューが終了した時点で、本実験の全体に対する主観評価を実施した。主観評価項目は、「リラックスする-緊張する」、「言いたいことが言えなかった」など計7項目であり、それぞれについて-3から3の7段階で回答を求めた。

4. 結果

4.1 発話量

まず発話量の定義であるが、書き起こした文字数や発話時間によって定義した場合、被験者の発話冗長性や発話速度に影響を受けてしまい、本来の目的としている発話量とはならない。

そこで本研究では、発話の内容によるいくつかのカテゴリを事前に設定し、実験によって得られた各被験者の発話を書き起こし、まずはこれらのカテゴリに分類した。次に、各分類の合計を算出し、これを各被験者の発話量として定義した。なお、発話の内容によるカテゴリは「Normanの7段階モデル²⁾(目標、意図、選択、実行、知覚、解釈、評価)」と、そこで分類しにくかった「評価2」と「探索」を新たに加えた9分類とした。各実験条件の発話量の平均値と標準偏差をFig.2に示す。

ここで、実験条件とタスク(タスク1はアイドリングタスクのため除き、タスク2,3を対象とする。以下、全ての統計処理において、タスク1を除くこととする)の二元配置分散分析の結果、実験条件による主効果が見られた($F(1, 14)=10.11$, $p<.05$)。但し、タスクの主効果、及び実験条件とタスクの交互作用は見られなかった($F(1, 14)=.60$, $p>.05$, $F(1, 14)=1.37$, $p>.05$)。

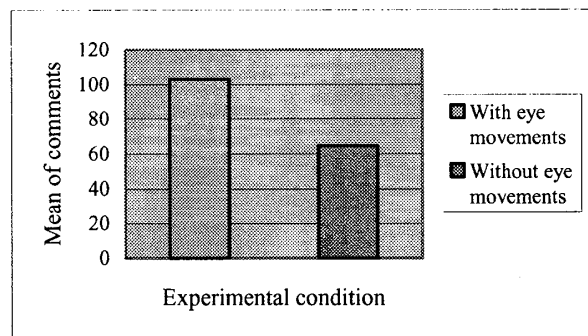


Fig.2 Mean of the quantity of comments

4.2 行動量

上述の分散分析により、実験条件による発話量の主効果が見られたが、行動量自体が視線あり条件で有意に多かった可能性もある。すなわち、行動量自体が多くそのために行動に関する発話量も多くなったという場合である。従って、行動量についても検証する必要がある。

行動量の定義は、レトросペクティブインタビュー法で使ったビデオから書き起こし、実験者2名により突合せを行った。ここで、実験条件とタスクの二元配置分散分析の結果、実験条件による主効果、タスクの主効果、及び実験条件とタスクの交互作用は見られなかった($F(1, 14)=2.10$, $p>.05$, $F(1, 14)=.59$, $p>.05$, $F(1, 14)=.06$, $p>.05$)。

4.3 主観評価

主観評価においては、すべての項目について実験条件、タスクに関する主効果、及び実験条件とタスクの交互作用について有意な差は見られなかった。

4.4 発話内容

次に、発話内容による各カテゴリでの平均値と標準偏差を Fig. 3 に示す。

ここで、実験条件とタスクの二元配置分散分析の結果、実験条件の主効果は、知覚、解釈、評価において見られた ($F(1, 14)=11.55, p<.05, F(1, 14)=13.39, p<.05, F(1, 14)=15.01, p<.05$) が、他の目標、意図、選択、実行、評価2、探索については見られなかった ($F(1, 14)=.78, p>.05, F(1, 14)=.56, p>.05, F(1, 14)=1.7, p>.05, F(1, 14)=1.7, p>.05, F(1, 14)=.01, p>.05, F(1, 14)=3.62, p>.05$)

また、タスクによる主効果は、評価のみで見られた ($F(1, 14)=4.85, p<.05$)。実験条件とタスクの交互作用については、全カテゴリとも見られなかった。

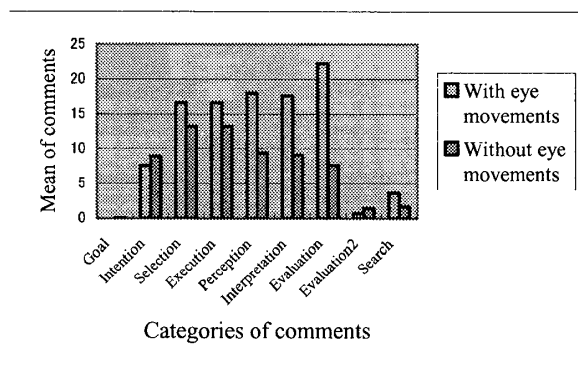


Fig.3 Mean of the quantity of comments for each category

5. 考察

5.1 実験条件による発話量

本実験では、ビデオ画像への視線情報の付加条件のみを実験条件とし、その他、全ての条件を同質にして実験を行い、この結果、視線あり条件において発話量は有意に多かった。行動量は、4.2 行動量の分散分析結果から、実験条件において有意に差があるとは言えない。これらの分析結果から、レトロスペクティブインタビュー時における視線軌跡の付加によって、被験者の発話量を有意に増加させたと言え、被験者の発話量の増加といった側面において、まずは本手法が有効であったと言える。

5.2 実験条件による発話内容

次に、実験条件による発話量の違いが、発話内容のどの部分に現れ、それが何を意味するのかを考察する。4.4 発話内容の結果から、知覚、解釈、評価において、視線あり条件で有意に発話量が多かったと言える。これらのカテゴリ

は、Norman の7段階モデルで言うところの評価過程に当たる。Norman は、実行までのカテゴリの実行過程と評価過程について、この2過程間に横たわるギャップを淵と称し、ここにインタフェースデザインの問題の源が存在し、いかに淵の橋渡しをするかがシステム・デザインの主要な課題であると主張している。そのためには、ユーザが、何を、どのように操作したかといった実行過程だけでなく、ユーザが何に対して、どのように解釈、評価したかといった評価過程が重要となり、本手法によって、この過程のユーザの発話を取得できることはシステム・デザインにおいて非常に有効であると考えられる。

更に、評価過程の発話内容に注目した場合、実験条件によってどのような違いの傾向があるかを検討する。書き起こした発話内容を一覧した結果、推測的な発話、例えば、「多分、この設定の組み合わせが良くないだろうなあ、と思い」といった表現が見られたため、評価過程に限って、推測的表現の数量を検討した。Fig. 4 に実験条件による推測的発話量の平均値と標準偏差を示す。

実験条件とタスクの二元配置分散分析の結果、実験条件の主効果が見られた ($F(1, 14)=17.35, p<.001$)。タスクの主効果、及び実験条件とタスクの交互作用は見られなかった ($F(1, 14)=.25, p>.05, F(1, 14)=.05, p>.05$)。すなわち、視線あり条件において推測に関する発話が有意に多いと言える。

実際に行った操作に関する言及は操作ログ解析などによっても取得することが可能であるが、これに加えてユーザの推測に関する発話抽出できるとにより、適切な解釈や誤解を含んだユーザの行動が本質的に理解できる。従って、本手法によって、インタラクション・システムの表面上の問題だけでなく、メンタルモデルに関連した問題を抽出できる可能性がある。

5.3 実験条件による主観評価

主観評価に関しては、実験条件での有意差は見られなかった。但し、両条件とも「言いたいことが言えた」と感じながら「緊張」状態にある一方で、視線なし条件では「緊張状

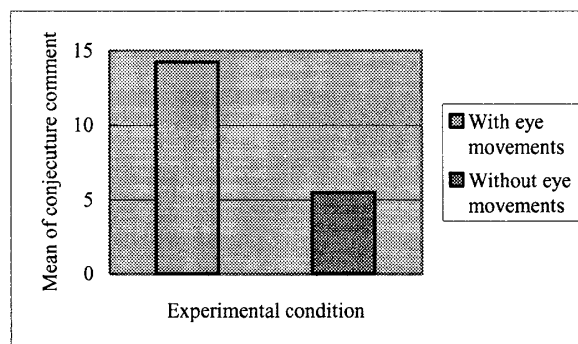


Fig.4 Mean of the quantity of conjecture comments during Evaluation process

態」が強い傾向が示唆された。今後も本評価手法による心的効果の側面も検討していく必要がある。

6. まとめ及び今後の課題

本研究の結果,以下の事柄が得られた。

- ・視線情報を用いたレトロスペクティブインタビュー法を提案し,プリンタドライバのユーザテストに本手法を適用した結果,以下の点から有効性が確認された。
- ・本手法によって,発話量が有意に増加した。
- ・発話内容は,評価過程に関する発話量が有意に増加した。
- ・視線あり条件評価過程で,推測的発言が有意に増加した。

今後は,発話内容のより詳細な分析,及び発話内容と行動の関係,他の生体情報を付加した手法などについて実験する必要がある。

参 考 文 献

- [1] Franca G., Maristella M., Paolo P. Model-based Heuristic Evaluation of Hypermedia Usability. in proceedings of the working conference on Advanced visual interfaces of ACM, 1998.
- [2] Mayhew D. J. Principles and Guidelines in Software User Interface Design. Prentice Hall, Englewood Cliffs, 1992.
- [3] Keith, K., Steve, E., Carnell, J. The Hunt for Usability:Tracking Eye Movements. in CHI '99 extended abstracts on Human factors in computer systems, 1999.
- [4] 海保博之,原田悦子,黒須正明:認知的インタフェース,p.72-75,新曜社(1995)
- [5] 海保博之,原田悦子:プロトコル分析入門,p.137-187,新曜社(2000)