

プロトコル解析手法による日本語WPの使い易さ評価

旭 敏之 神場 知成 宮井 均

日本電気(株) C&Cシステム研究所

各種情報処理機器の使い易さの評価手法として、「プロトコル解析技術」の応用を試みている。操作状況と発話内容の分析を基本とする本手法は、特に製品の改良に有効な評価手法であるが、今回は日本語ワープロ2機種を対象とした評価実験を行った。本実験により使い易さの面で種々の欠点が抽出されたが、それらを検討することにより日本語WPの評価手法としての本手法の適性を検討した。さらに、今後、実験/解析を支援するツールを作成していく予定であるが、その1つの可能性として、発話データに言語的処理を加えることにより、評価箇所を自動抽出するツールを考え、基礎検討を加えたので併せて報告する。

Usability Evaluation Of Japanese Word-Processors

By Protocol Analysis Method (in Japanese)

Toshiyuki ASAH, Tomonari KANBA, Hitoshi MIYAI

C & C Systems Research Laboratories, NEC Corp.

1-1 Miyazaki 4 cho-me, Miyamae-ku, Kawasaki, 213, Japan

“Protocol Analysis” method in psychology has been applied for evaluating usability of Man-Machine systems. This method is based on the observation of a user's behaviors and on the analysis of his verbal reports(protocol data). In experiments, two different types of Japanese word processors(JWP) were evaluated, and many problems on usability were extracted. In this paper, the results of the experiments are described, then suitability of this method for a JWP evaluation is discussed.

Also, some support tools for experiment and data analysis are under investigation. Primarily, an idea to pick up the check points from protocol data is discussed.

1. はじめに

人間との高度な対話機能を持つ情報処理機器の普及が、専門的なユーザに対してだけでなくオフィスや学校、一般家庭の間でも進んでいる。このような状況において、各種製品のユーザインタフェースの問題が一段と比重を増していることは誰もが認めることである。我々はユーザインタフェースの向上を目的とし、かつ実際の製品に効果的に反映させることを念頭においた、「製品の使いやすさの評価」に関する研究を進めている。すなわち「製品の使いやすさ」を実験的手法で評価し、その結果をフィードバックすることにより製品の改良を図る、あるいは次機種の設計時における検討データとすることが目的である。我々は既に、この実験手法として「プロトコル解析」の手法が有効であると考え、その効果を試行実験により確認してきた(1)。

一方情報処理機器のなかでも日本語ワープロの一般ユーザへの普及は、文書作成／保存のニーズの高さと、機体の低価格化とが相まって特にめざましいものがある。したがって、そのユーザ層も職業的な熟練者から中級者あるいは初心者にまで広がっており、さらに裾野を広げつつあるが、そのユーザインタフェースの問題は確たる指針が無いのが現状である。また、ワープロにおける作業は、ユーザの思考期間を除いて、ほぼ全てが種々のレベルの人間と機械の対話作業から成り立っている。以上のような事情から日本語ワープロにおいても、ユーザインタフェースの問題は特に重要である。さらに「製品評価」の側面から考えれば、ユーザとワープロとの対話そのものを効果的に評価する必要がある。

我々は日本語ワープロ2機種についてプロトコル解析の手法を応用して評価実験を行い、その効果等を検討したので報告する。さらに我々はこういった実験を通して、本評価手法をベースとして実験／解析手法の確立、あるいは自動化を目指した「製品評価システム」の構築を研究の最終目標としているが、今回はその端緒として検討した解析手順の一部自動化を図るツールの構想についても述べる。

2. 日本語ワープロの評価

日本語ワープロにおいて、重要な評価項目は次の2点に大別することができる。

- ① ディスプレイ、キーボード等のハードウェア的側面
- ② 対話手順、マニュアル等のソフトウェア的側面

①については例えばキーボードや入力方式について、主に作業効率(入力時間)を指標として多くの議論がある((2)など)。また、ディスプレイについても主観的な評価やあるいは作業による眼精疲労の研究など、これも多くの報告がなされている((3)など)。これらの評価手法の特徴としては人間工学的手法を用いて、定量的な検討を加えたものが多いことが挙げられる。しかしそういった評価手法を応用することが難しいためか、②のソフトウェア的な側面に関する評価の例は非常に少ない。

一方、日本語ワープロの特徴として、

- ④ 作業が対話的である
- ⑤ 初心者／中級者のユーザが多い
- ⑥ 機能が豊富である

などが挙げられる。初心者や中級者のユーザにとっては、文字の入力効率が多少変化することよりも、ワープロの備えている各機能の使いやすさの方が全体の作業効率に及ぼす影響は大きいと考えられる。したがって、ワープロの普及にともない初心者や中級者のユーザの増加が顕著であることも考慮にいと、日本語ワープロの使いやすさを効果的に評価するためには、②のソフトウェア的な側面の評価が不可欠であると考えられる。

さらに実際のワープロを用いた作業において、各機能の使いやすさはその使用状況によっても大きく影響を受けると考えられる。したがってワープロが備えている各機能は、実際の一連の作業状況下において評価するほうが望ましい。

我々は以上のような条件を考えた場合、プロトコル解析の手法が日本語ワープロの評価にも適していると考え、2機種の日本語ワープロを対象とした実験を行った。

3. プロトコル解析手法

プロトコル解析は被験者自身の口述内容をデータとするものである。その中でも「思考口述記録法 (Thinking Aloud Protocol Method)」は「被験者が実際に作業やある問題を解決している過程において、被験者自身に頭に浮かんだことを強制的に発話させ、これを記録/分析する」ことを基本としている。すなわち被験者の思考内容そのものの表現を試みる手法であることから、かなり高次の思考レベルのモデルに対しても有効な実験手法であるとされてきた。Ericsson ら(4)によると作業中の発話により得られるデータ(プロトコルデータ)は、短期記憶領域に一時的に保持された情報の一部であり、作業後のアンケートなどによるデータではその大部分が失われてしまう可能性があるとしている。したがって作業中に発話することにより、通常は忘れ去られてしまうような詳細な思考過程に関するデータを得ることができる。

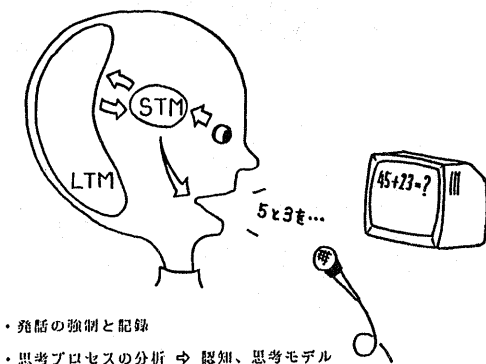


図1 プロトコル解析手法

4. 実験方法

心理学分野の実験手法であるプロトコル解析の手法を、日本語ワープロの使いやすさ評価に応用した。

被験者にはワープロ、マニュアル、タスクであるテキストを与えて、原則として自由に作業させる。
被験者には実験前の指示で作業中の発話を強制する。
 ただし、被験者が完全に操作に詰まってしまった場合などには、実験者が必要なアドバイスをこなう。

また、実験者は被験者の発話が途切れたときにも発話を促すべく発言するが、今回の実験ではほとんど必要がなかった。被験者の発話内容と操作状況はVTRに記録される。

解析はVTRに収録されたワープロ使用時の被験者の発話内容と操作状況を分析し、使いやすさの上の問題点の候補を抽出する。

実験は評価対象、被験者、タスクの組合せを変えて表1に示す7回を行った。

表1 実験設定の組合せ

実験	評価対象	被験者	タスク
実験1	ワープロ#1	中級者	タスク#1
実験2	ワープロ#1	中級者	タスク#2
実験3	ワープロ#1	中級者	タスク#3
実験4	ワープロ#2	初心者	タスク#1
実験5	ワープロ#2	中級者	タスク#1
実験6	ワープロ#2	初心者	タスク#2
実験7	ワープロ#2	中級者	タスク#2

<被験者>

各評価対象のワープロのユーザー層として予想される能力の被験者を用いた。

初心者：ワープロ経験2時間

中級者：ワープロ経験2.5年間

<タスク>

タスクはテキストである文書を、評価対象のワープロを用いて作成するように指示する。ただし、文書の保存や印刷などの操作は作業中に指示した。各タスクは、評価したい機能が必要となるように構成されている。例えば(タスク#2)の作業には以下の機能が必須である。

- ・文字入力(かな漢字、カタカナ、英数字、半角文字)
- ・中央揃え
- ・インデント
- ・TAB位置の設定
- ・他ファイルからの文書挿入
- ・印刷(印刷書式設定)
- ・文書ファイルの作成
- ・字体の変更

<記録>

- ・音声、全体像、手元（キーボード）を2台のVTRとカメラで記録した。
 - ・ワープロの画面変化も記録するため、画面出力信号を信号変換して3台目のVTRに入力した。
- 実験装置の設定を図2に示す。

<解析>

- ・収録されたVTRテープから発話内容を書き取る（プロトコルデータ）。
- ・プロトコルデータから、作業上の不都合を表現している箇所を抽出する。
- ・必要に応じて操作状況を再現し、使いやすさ上の欠点と思われる項目を抽出して行く。

5. 実験結果

得られたプロトコルデータから操作上の不都合を表現している箇所を抽出し、さらに使いやすさ上の問題点の候補となる項目を抽出した。各実験に要した時間（課題達成時間）と得られた評価項目の数を表2に示す。

表2 実験の所要時間

実験	時間（分）	評価項目
実験C-1	61	15
C-2	141	27
C-3	94	18
W-1	71	16
W-2	40	18
W-3	194	33
W-4	93	28

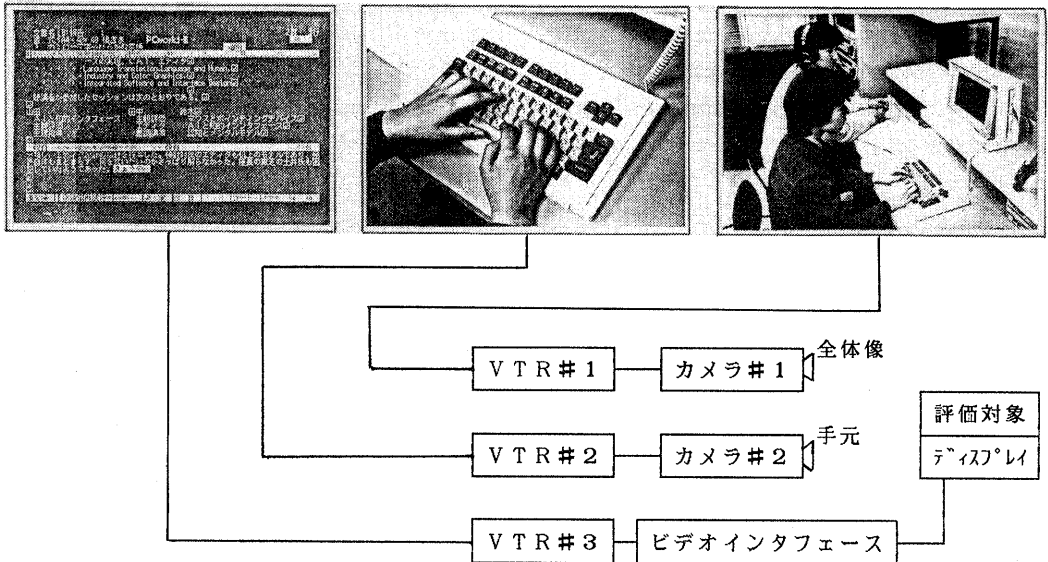


図2 実験装置の構成

13:38 記号で探した方が早いのかなあ。

13:47 あれ、これだけ？

14:13 芸が細かいから・・・。

14:21 あれー、でも飛ばし方がわからないや、カーソルの。

14:30 あっ、だめだ。せっかく見つかったのに。指定の仕方がわからない。
(カーソルを動かしたの?)

14:36 カーソルだって動かない。シフト矢印とかじゃないものね。
(矢印でいいかどうかわからない)

14:42 うん。困った。でも、こんなところで探したんじゃないかね、普通

14:51 ちょっと後回しにしよう。やってられない、こんなこと。

シフト+「機能」キーを押してメニューを見る
 野線 メニューの次ページを見る
 入 力 メニューの選択方法がわからない(カーソル移動による選択と思いついでいる)
 野線 入力をおさめ

図3 プロトコルデータ

図3に実験W-1で得られたプロトコルデータの例を示す。この例は中級者である被験者が罫線を入力しようとして、画面上に機能メニューを呼び出したが、その選択方法がわからず困惑している場面である。この機能メニューは機能キーを押すことで選択できるが、被験者はカーソルの移動による操作を試みていた。この事実はメニューの選択方法に関して、問題点の候補として指摘できる。

本実験の結果は、いずれもワープロの設計/開発部門にすべての問題点の候補を報告するという形で、フィードバックされた。

6. 考察

日本語ワープロ2機種についてプロトコル解析による使いやすさ評価実験の結果を報告した。特にワープロのソフトウェア的側面について、レベルを考慮した仮想的なユーザである被験者の振舞いと発話内容を検討することにより、実際の操作状況下での欠点の候補が多数抽出できた。

6.1 評価項目の分類

以下では得られた問題点の候補を分類し、本評価手法が日本語ワープロのこういった側面の評価に有効であったかを検討する。

操作上の問題（ユーザの予想との不一致）

- ・ 予想した機能がない
- ・ 予想した動作をしない
- ・ 予想外の動作をする
- ・ 待ち時間が長く不安になる
- ・ システムの応答に不自然さを感じる
- ・ 不親切に感じる機能
- ・ 目指す機能がないと思込む
- ・ ユーザが機能の解釈を間違え

操作上の問題（不明な/わかりにくい機能）

- ・ 操作がわからない
 - － 他操作で代行する
 - － マニュアルを読んでもわからない
 - － 操作の対象がわからない

- ・ 操作がわかりにくい
 - － 類推でわかる
 - － いくつか試みてからわかる
 - － マニュアルで発見する（容易/困難）
- ・ 操作を間違え
- ・ 操作がわずらわしい/イライラする

マニュアル、メッセージ、ヘルプ等の問題

（マニュアル）

- ・ 調べたい項目が発見できない
 - － 期待したページに載っていない
 - － どこを探せばいいかわからない
- ・ 機能不足（見出し、索引、関連事項の案内など）
- ・ 記述がわかりにくい/例、図がない

・ 量が多すぎる

（メッセージ）

- ・ メッセージの意味がわからない
- ・ メッセージを見逃す
- ・ 不安になるメッセージ

（（オンライン）ヘルプ）

- ・ 調べたい項目が発見できない
- ・ 記述がわからない/不親切

その他（画面/キーボードなど）

- ・ 画面が見づらい
- ・ キーボードが操作しにくい
- ・ キーを見逃す
- ・ キーが発見できない
- ・ 単純なキーの操作ミス
- ・ 機能の選択ミス
- ・ 同じ操作ミスを繰り返してしまう

以上の項目でわかるように、今回の実験により日本語ワープロの主にソフトウェア的側面に関して、総合的な評価項目が得られたことがわかる。しかも各評価項目は、使用状況やワープロの機能に関連して得られたものであり、実際に製品の改良を考えた場合にはより有効なデータであったと考えられる。

また、以上の項目は今回の実験に基づいたもので

あり、さらに実験を重ねることにより、充実／整理させていく予定である。

6.2 プロトコルデータの言語的処理

今回の実験では実験の解析（プロトコルデータの検討）に多大の時間（実験時間の3～4倍）を要した。我々は実験／解析手法の確立あるいは自動化を目的として研究を進めているが、今回は解析時において、プロトコルデータから操作上の問題点がある箇所を自動抽出するツールの実現の可能性を検討するため、発話内容の分類を行なったので以下に報告する。

（1）プロトコルデータからの抽出項目の分類

問題点を含む発話を自動的に抽出できれば、解析作業の効率化に効果が大きいと考えられる。そこで、プロトコル解析実験から得られた問題点を含む発話を、以下のように分類した。

- ①システムの動作に対するユーザの反応
 - a)自分の操作の間違いに気付いた時の発話
 - b)問題を解決した時の発話
 - c)システムのメッセージを理解できない時の発話
- ②ユーザの動作に伴う補助的な発話
 - a)操作方法がわからない時の発話
 - b)操作方法に自信がない（推測で行う）時の発話
- ③ユーザの感情を示す発話
 - a)不快な時の発話
 - b)不安な時の発話

次に、ワープロの評価実験のプロトコルデータをもとに、上記のそれぞれの項目を自動抽出するための手がかりとなる規則を検討する。

（2）システムの動作に対するユーザの反応

- ①自分の操作の間違いに気付いた時の発話パターンの分類
 - a)感動詞
 - b)望ましくない状況を示す単語＋完了を示す語尾
 - c)「～ちゃった、～てしまった」を含む表現
- a)として、この場合に特徴的なものとしては、「あ

れ（っ）？」がある（？は発音の際の昇調を示す）。

b)の「望ましくない状況を示す単語」には、対象領域に依存するものと、そうでないものがある。例えば、「違う、大変、だめ、やり直す」などは、どのような製品の評価の場合でも、望ましくない単語と考えることができるが、「消える、ずれる」などはワープロなどに限定されたものであろう。

c)の「～ちゃった、～てしまった」を含む表現は良くない状態を示すために用いられていることが多い。

②問題を解決した時の発話パターンの分類

- a)感動詞
 - b)望ましい状況を示す単語を含む
 - c)中立の状況を示す単語＋「（っ）てくれた」
- a)としてここで特徴的なものは、「そっかそっかそっか」「なるほどねー」などである。ただし、ここでは感動詞をやや広義に解釈した。
- b)の「望ましい状況を示す単語」として見られたものは、「わかる、良い、できる、合う」などである。
- c)の「中立の状況を示す単語」とは「やる、する」などであるが、これらに「てくれた」のような語尾が付いた表現は、良い意味に用いられている。
- ③システムのメッセージを理解できない時の発話
 - 「～って何？」「何よこれ」が見られた。

（3）ユーザの動作に伴う補助的な発話

- ①操作方法がわからない時の発話パターンの分類
 - a)感動詞
 - b)「わからない、困る、だめ、違う」などを含む
 - c)「どうして～？」
 - d)「できる」＋疑問（？） or 否定（ない）
 - e)操作を捜している時の発話
- a)として特徴的なものには、「えっ」「えー？」「んー」「あー」「あれ？」などがある。
- b)c)は、操作方法がわからないことを表現している。
- d)は、目的とする操作が、実は不可能ではないか、という疑念や確信を示している。発話例を示す。「できんのそんなこと」「～てできんのかな、できないね」
- e)は、何かの操作のためのキーを捜している時の発

話である。発話例を示す。

「～ってのはどうやんでしょね」「～てのなし？」

②操作方法に自信がない（推測で行う）時の発話パターンの分類

- a) 確認要求の表現（昇調を伴う）
- b) 確認要求の表現（疑問文の語尾であるもの）
- c) 「もし～たら～？」「もしかしたら～？」「もしかして～？」

一般に、操作方法に自信がない時は、操作の確認を横にいる実験者に求めるような発言が見られる。

a)は確認要求の表現であることが、昇調から判断されるものである。発話例を示す。

「～してもだめ？」「～とかしちゃって平気？」

b)は、確認要求の表現であることが語尾から判断できるもので、「～だね？～かな？」などがある。

c)は、その操作を試しにやってみる、という場合の表現である。発話例を示す。

「もしかして～なくちゃいけないんじゃない？」

(4) ユーザの感情を示す発話パターンの分類

①不快な時の発話

- a)感動詞
- b)望ましくない状況を示す単語
- c)望ましくない状況を示す慣用表現によるもの

a)としては、以下のようなものが見られる。

「あー」「えー」「クソー」

b)の例には、「不親切、困る、いや、遅い、いらいらする、不便、」などがある。

c)はb)と似ているが、特定の単語というよりも、慣用表現によるものである。発話例を示す。

「頭きた」「お話にならない」「気に入らない」

②不安な時の発話

これは、ワープロの評価実験では、最初にリセットをかける時と、最後に文書を保存して終了するときに限られていた。その時の発話例を示す。

「私それが心配なの。」「平気だね？」「ちょっと恐いね、この終了は、一気に押すのは。」

現在、以上の検討をもとに、発話の抽出を自動化する際の、単語辞書、構文解析などの問題点を検討中である。

ただし本分類は 6.1 の評価項目の分類とは無関係に検討されたものである。6.1 の評価項目を前提として本分類を再検討する、あるいは発話パターンの分類から評価項目を検討していくことが今後の課題の1つでもある。

7. おわりに

以上、プロトコル解析の手法をもちいて、日本語ワープロの使いやすさの評価実験を行なったので報告した。本実験手法はもともと心理学分野の手法であり、実験／解析の効率など、製品の評価手法としては検討の余地も多い。我々は今回報告したような評価実験を経験しながら、方式の確立と効率の向上を目的として研究を進めていく予定である。

参考文献

- [1] 旭、宮井
"プロトコル解析を応用した製品評価システム"
第34回情処大会論文集 PP 2181～2182 1987
- [2] D.A.NORMAN, D.FISHER
"Why alphabetic keyboards are not easy to use"
Human Factors vol.24 no.5 pp 509～519 1982
- [3] 鈴木 "OA機器と眼精疲労"
日本の眼科 vol.54 no.12 pp 1087～1095 1983
- [4] K.A.ERICSSON, H.A.SIMON
"Protocol Analysis"
The MIT Press, Cambridge, Mass 1984