

タスク実行の口頭指示による発話思考法を用いた
ユーザビリティ評価に関する研究

吉澤侑佑*1 堀 雅洋*2 加藤 隆*2

*1関西大学大学院 総合情報学研究科

*2関西大学 総合情報学部

目次

1. 背景
2. 評価実験
3. 実験結果
4. 考察
5. 今後の予定

▶ 2

発話思考法とは

- ▶ 課題を遂行する間に頭に浮かんだことを、すべて声に出して語る[Ericsson & Simon 1993]
 - ▶ 評価中の発話内容は対象の操作とともに記録・分析
- ▶ 事後アンケートからは得られない思考過程に関するデータが得られる

- ▶ 発話思考法の問題点
 - ▶ 操作と並行した発話の難しさ[Shuck 1966]
 - ▶ 頭に浮かんだ事をすべて声に出すのは容易ではない
 - ▶ ユーザビリティ評価に有用な発話を引き出すのは容易でない[Cooke 2010]

▶ 3

口頭指示を伴う発話思考法 [木原 2010]

- ▶ 評価者が実行したい操作を操作役(実験者)に指示し、操作役が操作



▶ 4

前実験の結果 [木原 2010]

- ▶ 口頭指示法によって
 - ▶ 発話のきっかけが与えられ、評価者の発話が促された
 - ▶ 完成度の低いプロトタイプでも、完成度の低さによる扱いにくさに評価者の関心が向くことを回避できた
 - ▶ プロトタイプと実機を比較
- ▶ さらに検討が必要な点
 - ▶ 評価者の発話内容についての検討
 - ▶ 同等の完成度を有する異なるユーザインタフェースの比較

▶ 5

研究目的

- ▶ 口頭指示法の有用性を検証
 - ▶ ユーザビリティ評価に有効な発話が得られるか
 - ▶ 発話内容から特定される問題点の抽出
 - ▶ 異なるユーザインタフェースにおける効果を比較
 - ▶ タッチ操作
 - ▶ マウス, キーボード操作

▶ 6

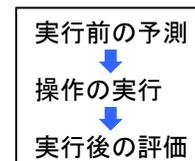
目次

1. 背景
2. 評価実験
3. 実験結果
4. 考察
5. 今後の予定

▶ 7

実験の概要

- ▶ 直接操作と口頭指示法で得られる発話内容の比較
 - ▶ 各操作段階における基本的な思考プロセス



- ▶ 評価データとして有用とされる「実行前の予測」と「実行後の評価」に関する発話に着目
- ▶ 「予測-実行-評価」の流れ(発話構造)を、発話の指針として評価者に提示

▶ 8

発話構造の提示

- ▶ 評価実験の事前説明で提示
 - ▶ 説明内容(口頭指示の場合)

操作時の注意点

以下の点に気を付けて操作してください

【操作前の予測】

その操作によって、どんな変化が起きると思うか声に出してください

【操作の実行】

操作内容は、声に出して操作役に伝えてください

【操作後の評価】

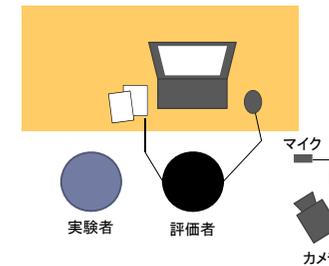
操作した結果はどうだったか声に出してください

直接操作の場合は
「操作を実行してください」と提示

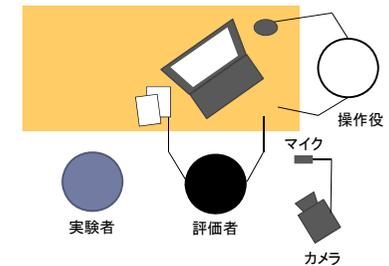
▶ 9

操作方法

直接操作



口頭指示



▶ 10

実験条件：操作方法（参加者間要因）

- ▶ 操作方法と発話構造提示の有無による条件分け
 - ▶ 条件A：直接操作・発話構造の提示なし
 - ▶ 従来の発話思考法に相当
 - ▶ 条件B：直接操作・発話構造の提示あり
 - ▶ 口頭指示を行わずに、発話構造の提示のみ行った場合の効果を検証
 - ▶ 条件C：口頭指示・発話構造の提示あり
 - ▶ 口頭指示と発話構造の提示によって全体の発話数、予測と評価に関する発話が最も多く得られると仮定

▶ 11

評価対象と評価協力者

- ▶ 評価対象(参加者内要因)
 - ▶ デジタルカメラ
 - ▶ タッチ操作(シングルタッチ, フリック, ドラッグ&ドロップ)
 - ▶ 写真管理ソフトウェア
 - ▶ マウス, キーボード操作
- ▶ 評価協力者
 - ▶ 総合情報学部3・4年生45名, 総合情報学研究科1年生3名
 - ▶ 男性30名, 女性18名
 - ▶ 発話思考法による評価経験なし

▶ 12

実験手順

▶ 事前説明と練習課題

- ▶ 実験者による発話思考法の実演
 - ▶ 条件B・C（発話構造の提示あり）では、発話の流れも説明
- ▶ 練習課題の実施

▶ 課題の実行

- ▶ 操作対象毎に練習課題（1問）と本番課題（1問）を実施
 - ▶ 順序効果はカウンターバランスをとって相殺
- ▶ 事後アンケート（7段階評価と自由記述）
 - ▶ 操作のしやすさ、発話のしやすさについて質問
- ▶ 評価者が5秒以上沈黙した場合の助言
 - ▶ 「思っていることをなんでも声に出して下さい」
 - ▶ 「今何を考えていますか？」

▶ 13

課題

▶ デジタルカメラ

- ▶ 練習課題
 - ▶ プレゼンテーションソフトによるプロトタイプをタッチパネル搭載PC上で操作
 - ▶ シングルタッチ、フリック、ドラッグ&ドロップを操作を練習
- ▶ 本番課題（ステップ数：13）
 - ▶ 撮影モードの変更、撮影、ファイルのコピー
 - ▶ シングルタッチ、フリック、ドラッグ&ドロップ操作を要する

▶ 写真管理ソフト

- ▶ 練習課題
 - ▶ メモ帳に文字を入力して保存
- ▶ 本番課題（ステップ数：12）
 - ▶ フォルダの作成、画像サイズの変更
 - ▶ いずれもクリックや文字入力などマウス・キーボード操作を要する

▶ 14

目次

1. 背景
2. 評価実験
3. 実験結果
4. 考察
5. 今後の予定

▶ 15

評価時間の比較

- ▶ 各課題のステップ数の差異をなくすため単一ステップあたりの評価時間（秒）に正規化
- ▶ 2要因（操作方法×評価対象）の分散分析

- ▶ 交互作用なし

		評価対象【有意差あり】		操作方法【有意差あり】	
				A.直接操作・構造なし	26.75
写真管理	デジカメ	30.77	43.23	B.直接操作・構造あり	32.84
		****	****	C.口頭指示・構造あり	51.41

* p<.05, ** p<.01, *** p<.005, **** p<.001

- ▶ 口頭指示の場合、操作内容を指示するため評価時間が長くなる
 - ▶ 前実験 [木原 2010] と同様の結果

▶ 16

発話数の比較

- ▶ 発話の時間的な区切り方
 - ▶ 0.5秒以上の沈黙または文の切れ目を目安に発話を区切る[海保 1993]
- ▶ 単一ステップあたりの発話数
 - ▶ 口頭指示の場合、実行に関する発話も含まれているため、実行に関する発話を除く
- ▶ 2要因(操作方法×評価対象)の分散分析
 - ▶ 交互作用なし

評価対象【有意差あり】		操作方法【有意差あり】	
写真管理	デジカメ	A. 直接操作・構造なし	B. 直接操作・構造あり
5.35	7.56	5.43	5.77
<		*	
		C. 口頭指示・構造あり	
		8.16	

* p<.05, ** p<.01, *** p<.005, **** p<.001

- ▶ 口頭指示法を用いることで発話数が増加
 - ▶ 前実験[木原 2010]と同様の結果

発話内容を考慮した区切り方 [Cooke2010]

- ▶ 意味的なまとまりを意識した区切り方
 - ▶ 発話の単位は、文、節、句、単語、感嘆詞とする
 - ▶ 意味的なまとまりを意識して発話を区切る
 - ▶ 2秒以上の沈黙がある場合は発話を区切る

▶ 発話数

操作方法	発話数
A. 直接操作・構造なし (n=16)	1975
B. 直接操作・構造あり (n=16)	2274
C. 口頭指示・構造あり (n=16)	3176

発話内容の分類

▶ 4種類の分類項目

項目	内容
予測	評価者の操作結果に対する予測
実行	評価者が評価対象を操作した際の発話
評価	評価者の操作結果に対する評価
その他	上記の分類に属さない発話(操作マニュアルやタスク内容を読んでいる際の発話を含む)

▶ 分類結果(2名の評定者による平均)

操作方法	予測	実行	評価	その他
A. 直接操作・構造なし	119.5	371.5	498.5	985.5
B. 直接操作・構造あり	318.5	381.0	661.0	913.5
C. 口頭指示・構造あり	472.0	838.5	809.5	1056.0

「予測」と「評価」に関する発話数の比較

- ▶ 単一ステップあたりの予測と評価の発話数
- ▶ 2要因(操作方法×評価対象)の分散分析
 - ▶ 交互作用なし

評価対象に有意差あり		操作方法に有意差あり	
写真管理	デジカメ	A. 直接操作・構造なし	B. 直接操作・構造あり
1.75	3.0	1.55	2.45
<		****	
		C. 口頭指示・構造あり	
		3.20	

* p<.05, ** p<.01, *** p<.005, **** p<.001

- ▶ 普段使い慣れていないデジカメの場合、迷いが多く生じて発話数が増加
- ▶ 発話構造を提示することで、予測と評価の発話数が増加
- ▶ 口頭指示法を用いることで、予測と評価の発話数が増加

評定者による分類結果の相違

- ▶ 発話内容の分類は2名の評定者で行った
 - ▶ 2名とも全評価者(48名分)の分類を行った
 - ▶ 互いに相談せず各自で分類を行った

操作方法	評定者	予測	実行	評価	その他
A (n=16)	1	96	579	542	758
	2	143	164	455	1213
B (n=16)	1	327	488	719	740
	2	310	274	603	1087
C (n=16)	1	437	826	879	1034
	2	507	851	740	1078

▶ 21

発話内容の分類に関する評定者間信頼度

- ▶ 評定者間の分類結果の相違を比較
 - ▶ 単純一致度(%)
 - ▶ 偶然一致したものを考慮していない
 - ▶ Krippendorffの α 値[Freelon 2010]
 - ▶ 偶然一致したものを考慮

▶ 22

評定者間信頼度

- ▶ 単純一致度(%)
 - ▶ 1要因の分散分析
 - ▶ 交互作用なし
- ▶ Krippendorffの α 値
 - ▶ 1要因の分散分析
 - ▶ 交互作用なし

****	A. 直接操作・構造なし 61.8 B. 直接操作・構造あり 73.3 C. 口頭指示・構造あり 81.4	****	A. 直接操作・構造なし 0.41 B. 直接操作・構造あり 0.62 C. 口頭指示・構造あり 0.74	****
*		****		****
		****		****

* p<.05, ** p<.01, *** p<.005, **** p<.001

- ▶ Krippendorffの α 値は0.7以上が望ましいとされる
 - ▶ 条件A, Bにおいて実行に関する分類の相違が顕著
 - ▶ 口頭指示(条件C)の場合, 評定者の相違による影響を受けにくい

▶ 23

事後アンケート

- ▶ 発話のしやすさと操作のしやすさを7段階評価
 - ▶ 2要因(操作方法×評価対象)の分散分析
- ▶ 発話のしやすさ【有意差なし】
 - ▶ 口頭指示による影響は少ないと考えられる
- ▶ 操作のしやすさ
 - ▶ 交互作用あり
 - ▶ 写真管理ソフトについて
 - ▶ 条件A-C, 条件B-Cに有意差あり(p<.05)
 - ▶ マウス, キーボードは普段から使い慣れているため, 口頭による指示が煩わしいと感じられる

▶ 24

問題検出件数

- ▶ 評価者の発話から評価対象の問題点を検出
 - ▶ 評価者一人当たりの問題点検出件数
 - ▶ 2要因(操作方法×評価対象)の分散分析
 - ▶ 交互作用なし

評価対象に有意差あり		操作方法に有意差あり	
写真管理	デジカメ	A.直接操作・構造なし	B.直接操作・構造あり
1.89	< 3.63	3.88	5.75
****		*	

* p<.05, ** p<.01, *** p<.005, **** p<.001

C.口頭指示・構造あり
6.75

- ▶ 口頭指示と発話構造を提示することで問題検出件数が増加

▶ 25

目次

1. 背景
2. 評価実験
3. 実験結果
4. 考察
5. 今後の予定

▶ 26

考察 1/2

- ▶ 異なるユーザインタフェースにおける効果

写真管理ソフトウェア (マウス・キーボード操作)	<	デジタルカメラ (タッチ操作)
-----------------------------	---	--------------------

※有意差のあった項目

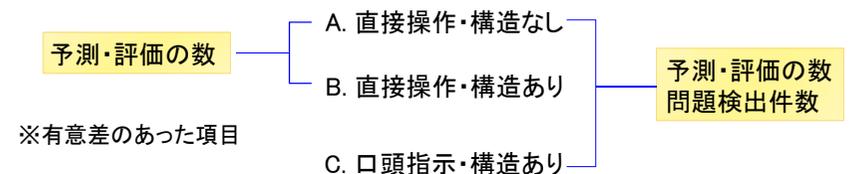
評価時間
発話数
予測・評価の数
問題検出件数

- ▶ タッチ操作
 - ▶ 全体の発話量, 予測と評価に関する発話数が増加
 - ▶ 普段使い慣れていないため, 迷いが多く生じたと考えられる
 - ▶ 問題点検出件数
 - ▶ 操作の行為(タッチ操作)に関する問題点も検出された

▶ 27

考察 2/2

- ▶ ユーザビリティ評価に有効な発話が得られたか



- ▶ 予測・評価の数
 - ▶ 発話構造の提示のみで有意に増加
- ▶ 問題検出件数
 - ▶ 発話構造の提示だけでは有意差なし
 - ▶ 口頭指示と発話構造の提示を行うことで有意に増加

▶ 28

今後の予定

- ▶ **問題点検出についてより詳細な検討**
 - ▶ より詳細な分類項目に基づく発話分類と問題点検出
 - ▶ 本発表では「予測-実行-評価」の流れに基づく分類
 - ▶ Normanの行為の7段階モデルに基づいた分類を用いる
 - ▶ 評定者間信頼度の検討
- ▶ **評価者への負担の軽減**
 - ▶ 操作の指示を容易にするための事前説明や指示方法の教示を検討
- ▶ **実施・分析手順のマニュアル化**
 - ▶ Webサイト評価の実案件への適用