

## 3. ユーザテストによるエラーの検討

### 3.1 本章の目的

本章では、逐次音声対話方式を用いて市販化されたカーナビを対象に、ユーザビリティという観点から、対話中の音声呈示においてどのような問題が潜んでいるかを探る。これにより理解しやすい音声呈示法を検討する糸口となるはずである。この点において、既存の研究において問題点を示した例はほとんど見当たらない。そこで、本章では、ユーザテストという考え方を導入し、対話エラーにおける音声呈示の問題点について検討した。

### 3.2 ストーリーボード評価システム

#### 3.2.1 逐次操作型 UI

本研究で対象とする逐次操作型の UI は操作に開始と終了があり、逐次的に操作が進められ、同時にある程度の可逆性をもっている(図 3.1)。また、入力するコマンドによって別の操作系に分岐する箇所があり、階層構造をもっていることが特徴である。

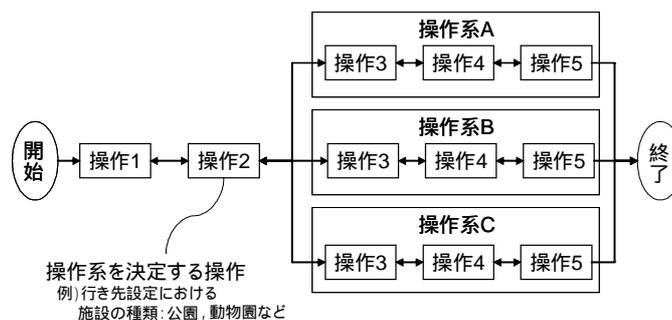


図 3.1 逐次操作型 UI の特徴

#### 3.2.2 ユーザテスト手法

ユーザビリティの評価法にはさまざまなものがあるが、特に製品を対象とした場合、その UI にどのような問題があるのかを知りたいときは、ユーザテストを行うことで、ユーザがある目的を達成するために製品を一つの道具や手段として使うときの問題点が明らかになる [3-1]。ユーザテストにはさまざまな方法がある。例えば Keystroke Level Model (KLM) や Dual Task Model (DTM) といった方法があげられる [3-2][3-3]。

KLM は操作の手数とユーザの認知的負荷を操作の所要時間に換算して評価する手法である。DTM はナビゲーションや自動車電話のように運転操作をしながら同時に操作されるような機器の操作性能を評価することができる。このモデルは、人間の目や手といった入出力器の動作時間とともに、機器の処理や表示にかかる時間を記述することができる（図 3.2）。これらの手法の特徴としては、評価の構造を探ることを目的として評価項目間における定量的な距離を測ることがあげられる。

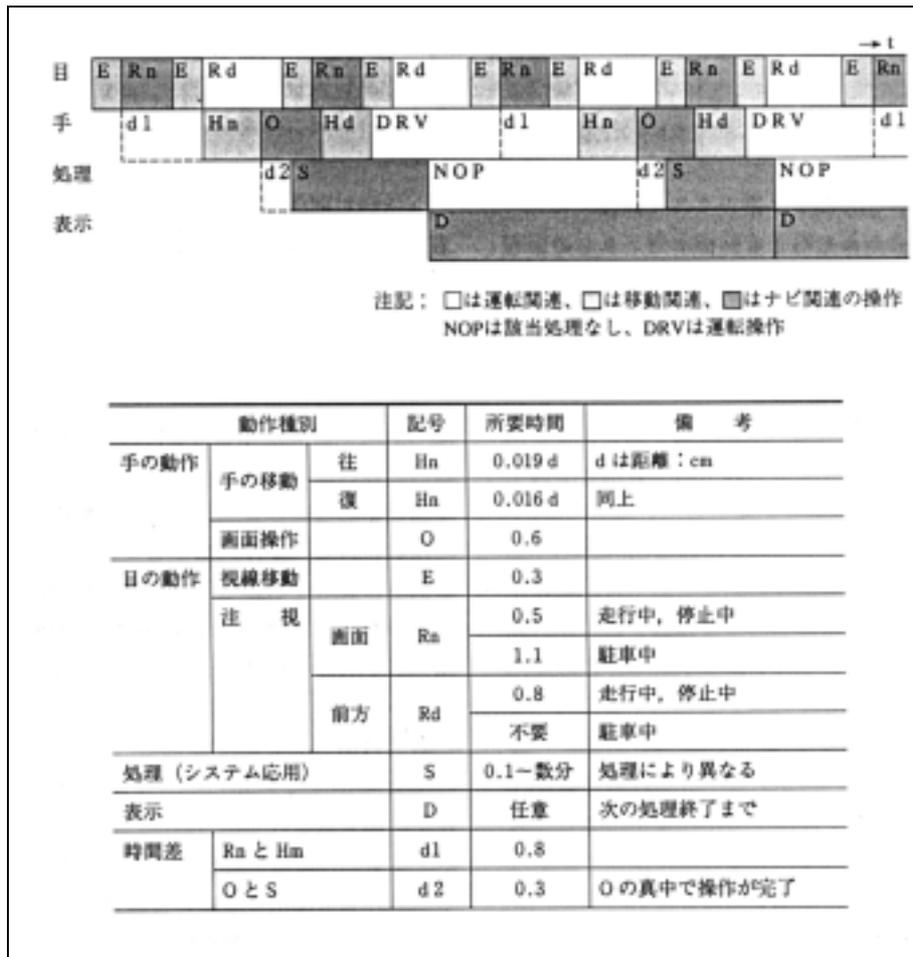


図 3.2 DTM の構成 (初心者, 走行中, 停止中)

(参考文献[3-2]pp.136 より転載)

一方、ユーザのメンタルモデルなど、評価における心的な関係を見たい場合は観察法やプロトコル分析法が適している [3-4]。本章では、ユーザの行動パターンをみることで、機器操作に対してどのようなメンタルモデルをもっているかを検討することで、以降の章において実験室環境を構築する材料としたいと考えた。そこで、観察法を用いてユーザの行動パターンを調べ、問題点を抽出する形をとった。

ここで、車載情報機器に実装されるソフトウェアは、機能の増加に伴って階層構造が複

雑になってくることが予想されるので、可逆的な階層操作におけるユーザの状態遷移を記述することは重要といえる。しかし、従来の観察法は情報を直列的に記述するものがほとんどのように見受けられる。そこで、エラー率の計算と結果に至るまでの行動パターンを詳細に分析するために、階層構造をもつソフトウェア使用時の状態遷移を記述することができる評価システムを構築した。

具体的には観察システムを構築し、対話作業中にユーザの立場から見てどのようなエラーが発生しているかをビデオに映像・音響収録して検討した [3-5][3-6]。

### 3.2.3 観察法とプロトコル分析法を用いた行動観察システム

観察法は、統制された実験環境において、人間がどのような行動をとったか、どのようなエラーを起こしたかを実験者による観察を通して同定する手法である [3-4]。この手法は観察者としての経験がある程度必要となるが、直接的に問題点を明らかにできる点において有用な方法である。プロトコル分析法はユーザテストの一手法で、被験者が与えられた何らかの課題を解決しようとしている場面で、推理、判断や思考のプロセスを口に出して述べてもらい、それを記録し、分析する方法である [3-7]。観察法もプロトコル分析法も、記録はビデオカメラやマイクを用いて行うことで再現性を持たせる。図 3.3 では、2 台のビデオカメラを用いている。一方のカメラはキーボード操作の様子、被験者の表情ならびに発話内容の音声を収録している。もう一方のカメラはモニタ画面の様子を記録している。



図 3.3 プロトコル分析のための実験環境 (参考文献 [3-7] p.159 より転載)

本研究においては、この 2 つの手法を応用しソフトウェアにおける階層操作の様子を観察できるストーリーボード評価システムを考案した。ストーリーボードとは、対話型イン

タフェースを設計する際に用いられ、一覧性をもって可逆的な操作フローを明確に記述することのできる画面遷移の記述形式である[3-4] (図 3.4). 図に示したストーリーボードは GUI 画面設計の例であるが、左から右に逐次的な遷移を、上から下に操作系(モード)の遷移を記述している。これによって、本来は一度にひとつの画面しか見られないところを相互の関連性も含めてユーザの操作順序という観点から概観することができる。これを観察法に組み込むことによって、ユーザの操作履歴を詳細に記録することができると考えられる。ストーリーボード評価システムは 1) 車内 HMI 収録システムと 2) ストーリーボード評価シートから構成される。

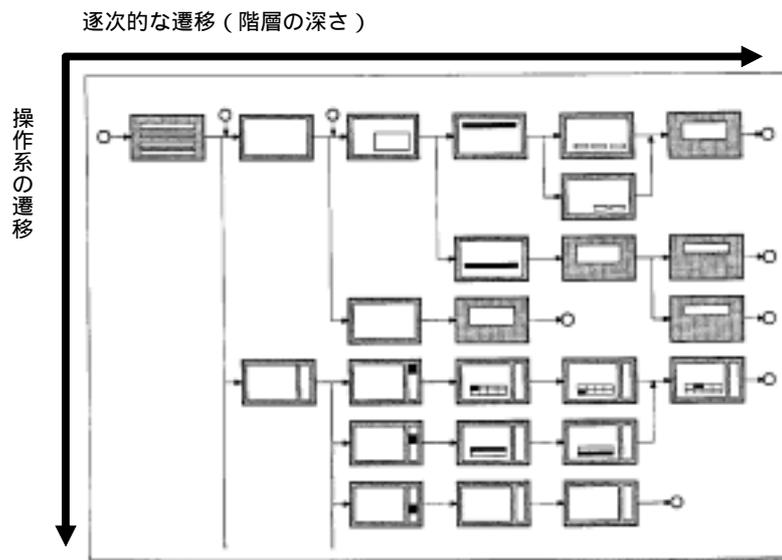


図 3.4 ストーリーボードによる状態遷移の表現(参考文献[3-6]p.210 の図に加筆)

### 3.2.4 車内 HMI 収録システム

車載情報機器を使用しながらの運転環境は、マルチタスク HMI である。このような HMI における対話の状況をさまざまな角度から収録・測定を行うために、映像、音響、ならびに生理データを同時に収集できるシステムを構築した。評価に用いる収録システム構成を図 3.5 に示す。

まず、助手席側に設置された CCD カメラ(カメラ)により、被験者の作業の全体的な様子および、表情の動向を収録する。同時に、後席に設置されたデジタルビデオ(DV)カメラ(カメラ)により、音声入力ボタンの操作状況および、モニタ画面の様子を収録する。

音響は、ダミーヘッドマイクを用いて被験者の発話内容や操作音、およびカーナビの音声ガイドを立体収録した。これは、立体音響を用いることにより、作業空間の音響情報をより高い臨場感をもって捉えることができる[3-8]ため、効果的に解析作業を進めることができると考えられる。

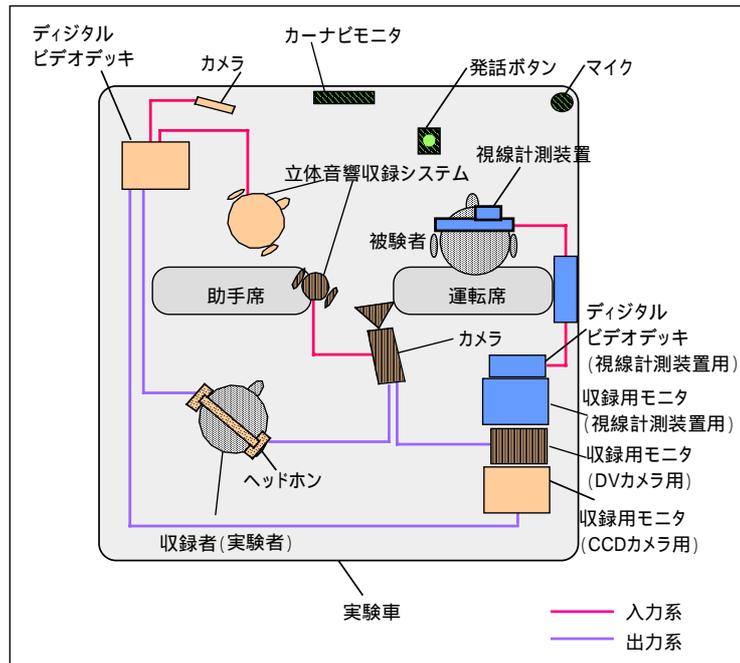


図 3.5 CCD カメラ設置と立体音響収録システム

### 3.2.5 評価システム

評価システムは(1)2画面構成ビデオモニタリングシステムと,(2)ストーリーボード評価シートから構成される。

#### (1) 2画面ビデオモニタリングシステム

2画面ビデオモニタリングシステムは,2台のS-VHSビデオデッキをタイムコードで同期させ,アナログビデオエフェクタを通して2画面構成に編集したものを,画面表示する仕組みになっている(図3.6),これにより,被験者の動向,プロトコルを2画面構成で観察することができる(図3.7).

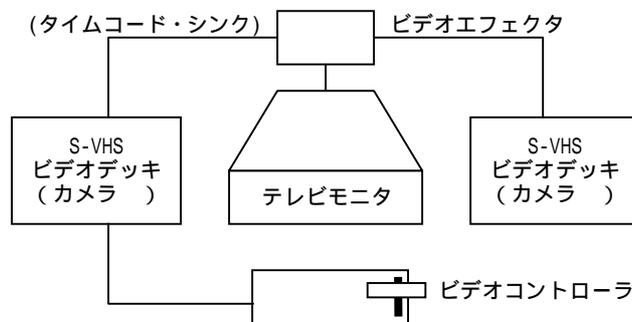


図 3.6 2画面構成ビデオモニタリングシステムの構成



図 3.7 2画面構成画面

## (2) ストーリーボード評価シート

ストーリーボード評価シートは、(2-1) ストーリーボード記述部、(2-2) 被験者の動向記述部からなる。この記述シートを構成する各部は、対象とする人間-機械系を反映するように設計される。本研究では、音声認識型カーナビの人間-機械系を反映する内容の評価シートを作成した(図3.8)。

### (2-1) ストーリーボード記述部

ストーリーボード部は、正しいフローを記述するA) システムセクションと、被験者の行ったフローを記述するB) ユーザセクションから構成される。

#### A) システムセクション

ここでは、カーナビのモニタ画面に表示される画面が表される。モニタ画面図はそれぞれ、正しい流れに沿ったイベントを実行する際の、操作フローをストーリーボードとして記述する。同時に、ガイド音声、復唱音声、発話ボタン、および、被験者が発話することになっているコマンドのフローを記述する。

#### B) ユーザセクション

ここでは、実際に被験者が行った行動のフローがシステムセクションと同様に記述される。記述は主として、どのような流れでコミュニケーションにエラーが起きたかについて視覚的に行われる。

### (2-2) 被験者の動向記述部

ここでは、被験者の動向を記述する。具体的には、視線の遷移、および被験者の自由発話内容を記述する。

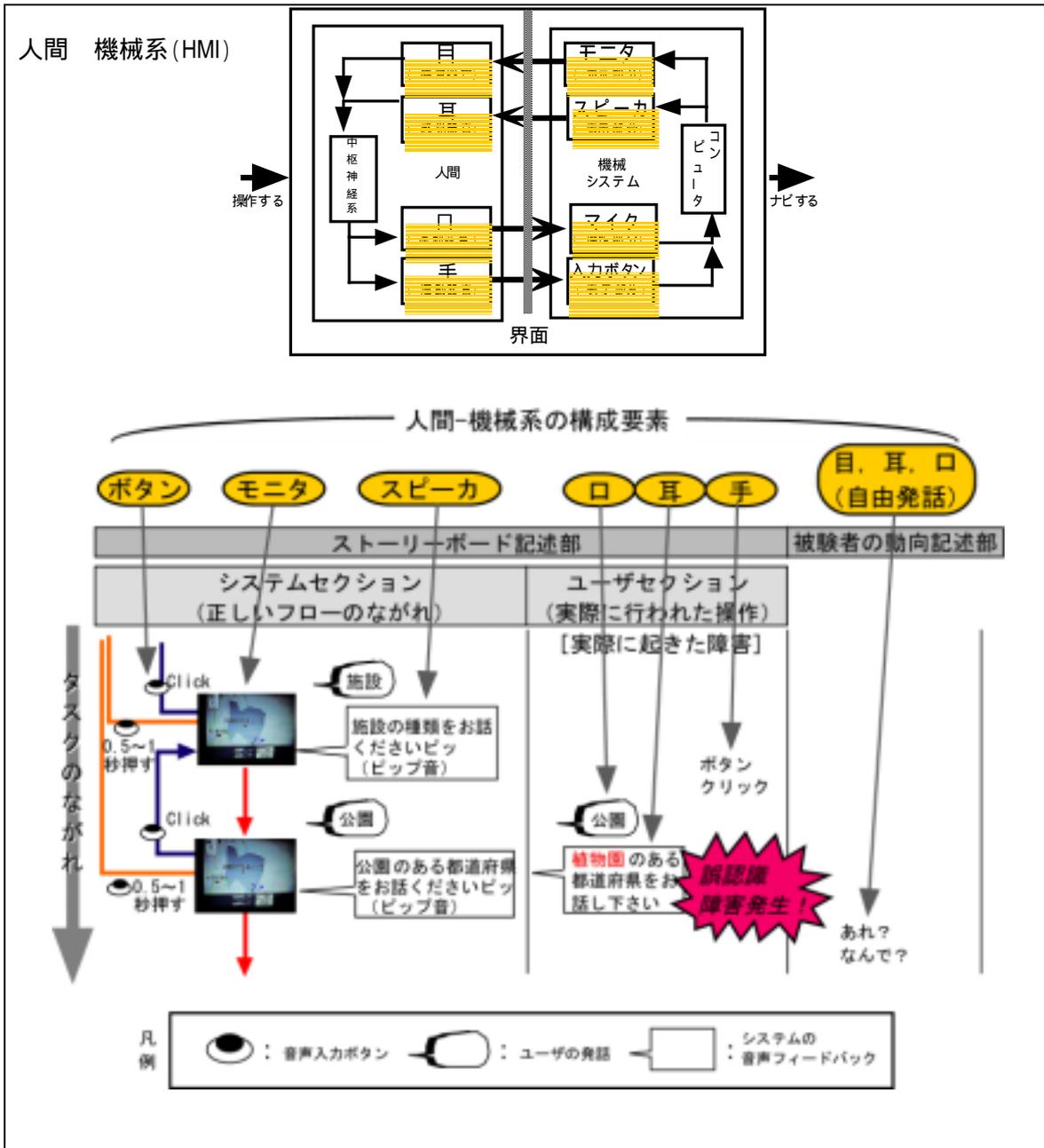


図 3.8 人間-機械系とストーリーボード評価シートとの対応

### 3.3 実験方法

#### 3.3.1 実験システム

使用した車載情報機器は、市販の逐次音声対話型カーナビ3機種であった。いずれも基本的な操作はガイダンス呈示 発話ボタン押下 発話 音声認識 次のガイダンスというながれになっている(図3.9)。

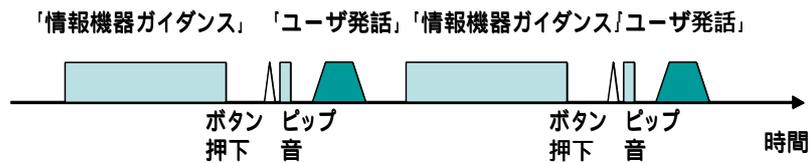


図 3.9 基本操作の流れ

### 3.3.2 実験課題

実験課題は、1) 目的地の住所による検索、2) 目的地の電話番号による検索、3) 周辺情報の検索、ならびに 4) 地図の表示形式の変更であった。コマンドの長さは平均 5.4 音節であった。実験はカーナビの操作のみを観察するため、駐車状態で行った。

### 3.3.3 評価項目

評価項目はストーリーボード評価システムを用いた観察法とプロトコル分析法によるエラー抽出と行動のパタン化であった。

## 3.4 実験手続き

### 3.4.1 実験場所および被験者

早稲田大学人間科学部南門駐車場において行った。被験者は、年齢 20 才から 24 才までの健康な男子 1 名、女子 3 名、計 4 名で、いずれもカーナビ操作に関しては初心者であった。

### 3.4.2 実験手順

以下の手順に従って実験を行った

#### 手順 1

被験者に、実験に関係のある部分のカーナビ操作マニュアルを渡し、前もって読んでもらった。

#### 手順 2

実験を始める前に、実験に関して教示を行った。教示の中で、操作中に感じたこと、考えたことについて、実験課題遂行にあたりできるだけ発話するように指示した。

また、この時点でカーナビの基本操作として音声入力操作と入力結果の取消操作についてマニュアルを見なくても操作可能になるまで学習してもらった。

#### 手順 3

実験課題を操作してもらった。入力コマンドは必要に応じてマニュアルを参照してもらった。実験の様子を図 3.10 に示す。

#### 手順 4

実験終了後、被験者にカーナビに関する質問紙に記入してもらった。また、インタビュー形式にて実験に関する内省報告を採取した。

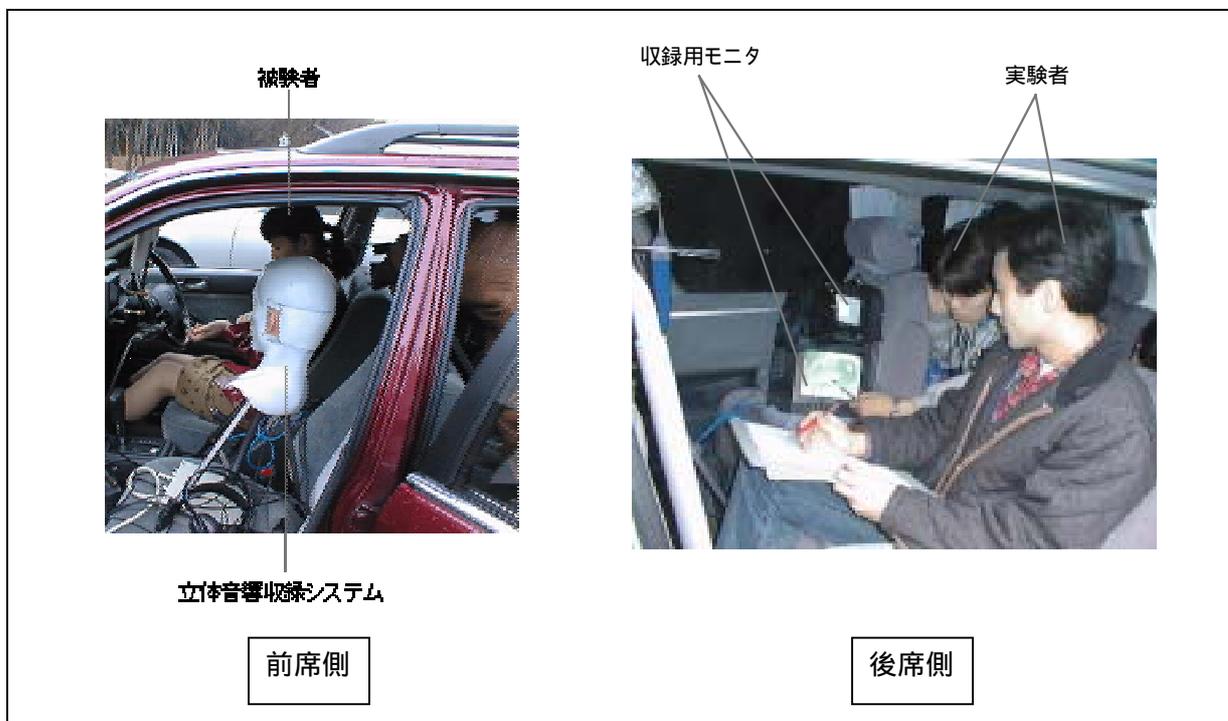


図 3.10 実験風景

### 3.5 結果

#### 3.5.1 エラーの分類

分析の結果、以下の 4 パタンのエラーが観察された。具体的には以下のとおりである。

##### パターン 1)「発話タイミングエラー」

これは、ユーザの発話タイミングに関するエラーである。具体的には、カーナビが音声の認識を開始する前に被験者が発話してしまったケースと、発話を行なわないうちにコマンド認識の制限時間が終了してしまったケースが認められた。図 3.11 は、タスク 1)において被験者が、マシン側からの発話許可のピップ音が呈示される前に発話してしまった結果、認識されなかった例である。「神奈川」と発話した時点で発話許可のピップ音が呈示されている。

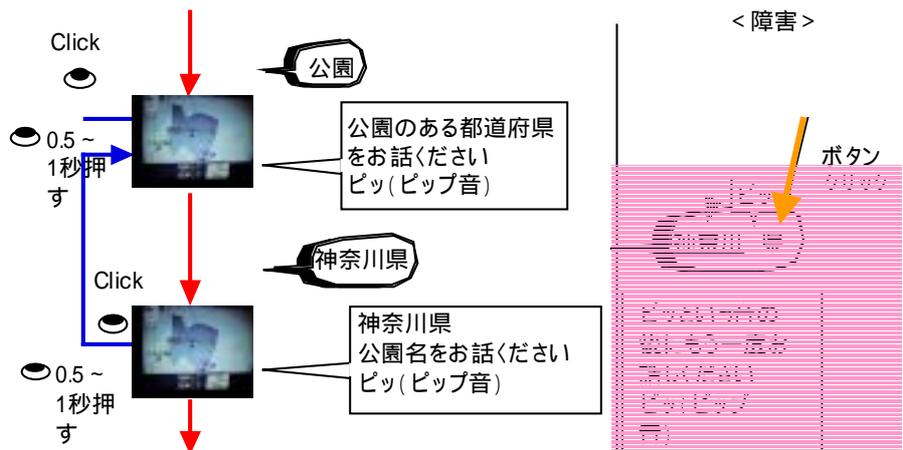


図 3.11 「発話タイミングエラー」

### パターン 2) 「誤認識エラー」

これは、ユーザは正しいタイミングでデータベース上に存在するコマンドを発話したにもかかわらず、カーナビが異なったコマンドとして認識を行ってしまい、目的のタスクが達成できなかったエラーである。図 3.12 は、タスク 1) において被験者が正しい手順で「公園」と発話したにもかかわらず、マシン側が「植物園」と誤認識してしまったものである。

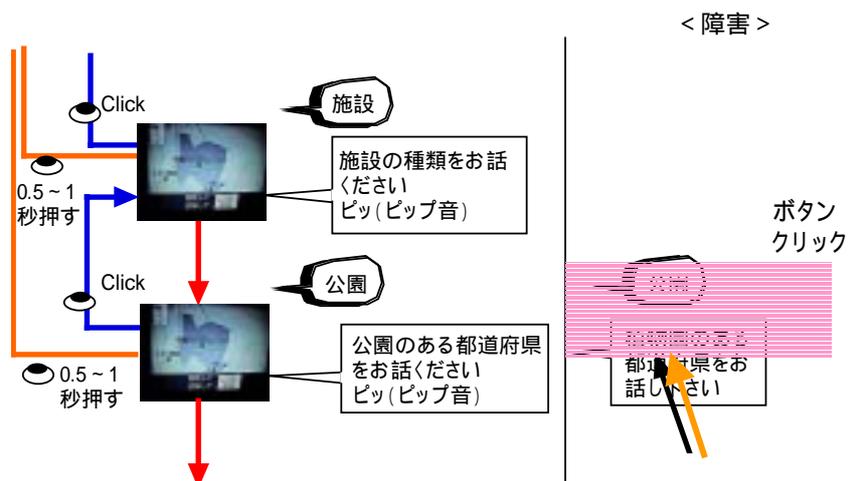


図 3.12 「誤認識エラー」

### パターン 3) 「不認識エラー」

これは、ユーザは正しいタイミングでデータベース上に存在するコマンドを発話したにもかかわらず、カーナビがそのコマンドを認識しなかったというエラーである。図 3.13 は、タスク 2) において被験者が「ハチ (8)」というコマンドを正しい手順で発話したにもかかわらず、マシン側が認識しなかったというものである。

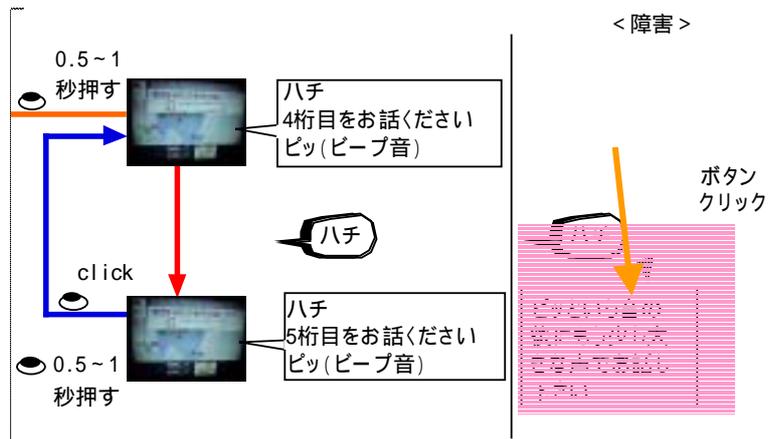


図 3.13 「不認識エラー」

#### ボタン 4) 「誤操作エラー」

これは、ユーザー側の誤った操作が原因で発生したエラーである。具体的には、ユーザが発話ボタンを押さずに音声コマンドを発話してしまったために発生したエラーが挙げられる。図 3.14 は、タスク 1) において被験者が発話ボタンを押さずに「東京都」と発話してしまった結果、認識されなかったというものである。

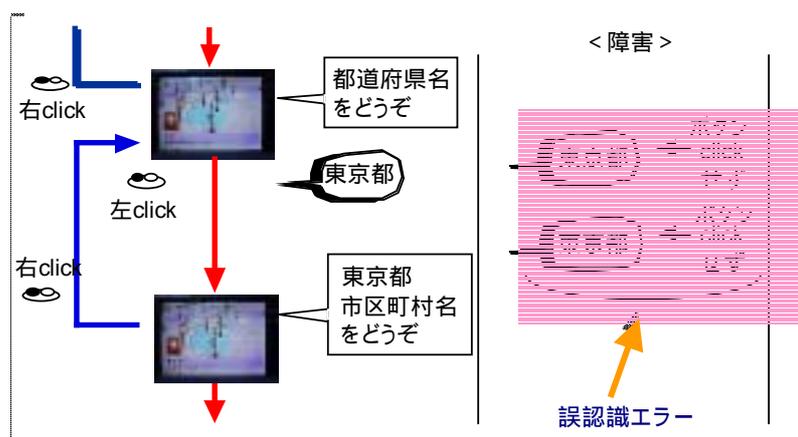


図 3.14 「誤操作エラー」

### 3.5.2 エラーの発生割合

次に、観察された各エラーの発生割合を求めたところ、表 3.1 に示すとおりとなった。

表 3.1 エラー発生の割合

	誤認識エラー	不認識エラー	発話タイミングエラー	誤操作エラー
A社	33%	35%	19%	13%
B社	29%	48%	10%	13%
C社	41%	13%	33%	13%

### 3.5.3 エラーのループ

誤認識エラーの発生後に、連続して複数回の誤認識エラーを起こす、誤認識エラーのループが起きるケースが、最初に誤認識エラーが発生した直後に平均 60%の割合で発生が観察された。さらに、一旦ループが起きるとその 60%が正しい操作系に復帰することなくタスク失敗を起こしていることが明らかとなった。この現象について操作イベントの発生と関連付けをおこなって分析を行った結果、図 3.15 に示すような共通のメカニズムが存在することがわかった。

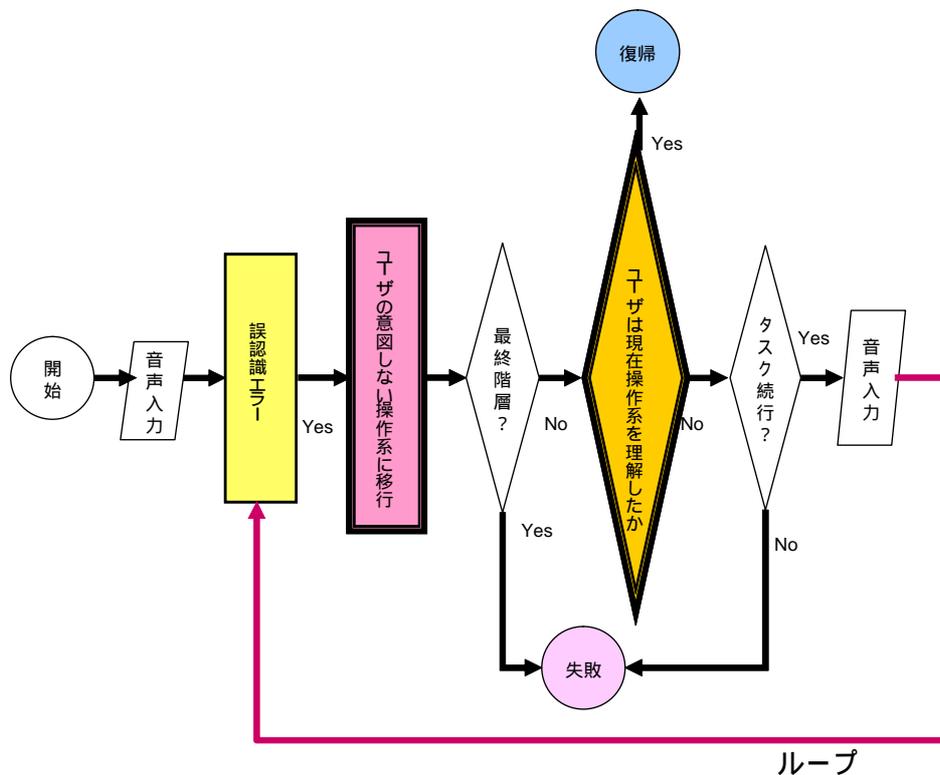


図 3.15 ループの発生フロー

これは、トリガとなる誤認識エラーが発生した際に、目的とする操作系から逸脱する内容であることにユーザが気づかないと、目的の操作系で発話すべきコマンドを逸脱した操

作系において発話してしまうという行動が見られる。その結果、さらに誤認識エラーを発生させ、ユーザは現在操作系を見失ってしまい、操作完遂に失敗してしまう結果を招いてしまうメカニズムとなっている。

図 3.16 で示した実際に観察された例では、被験者は山下公園に行き先設定をしたいが、施設の種類の入力結果が誤認識されてしまい、「植物園」を検索する操作系に入ってしまった。しかしその音声ガイダンスを理解できず、用意していたコマンド「山下公園」を発話すると、データベース上にないコマンドであったため、「江ノ島植物園」とさらに誤認識エラーが発生してしまっている。

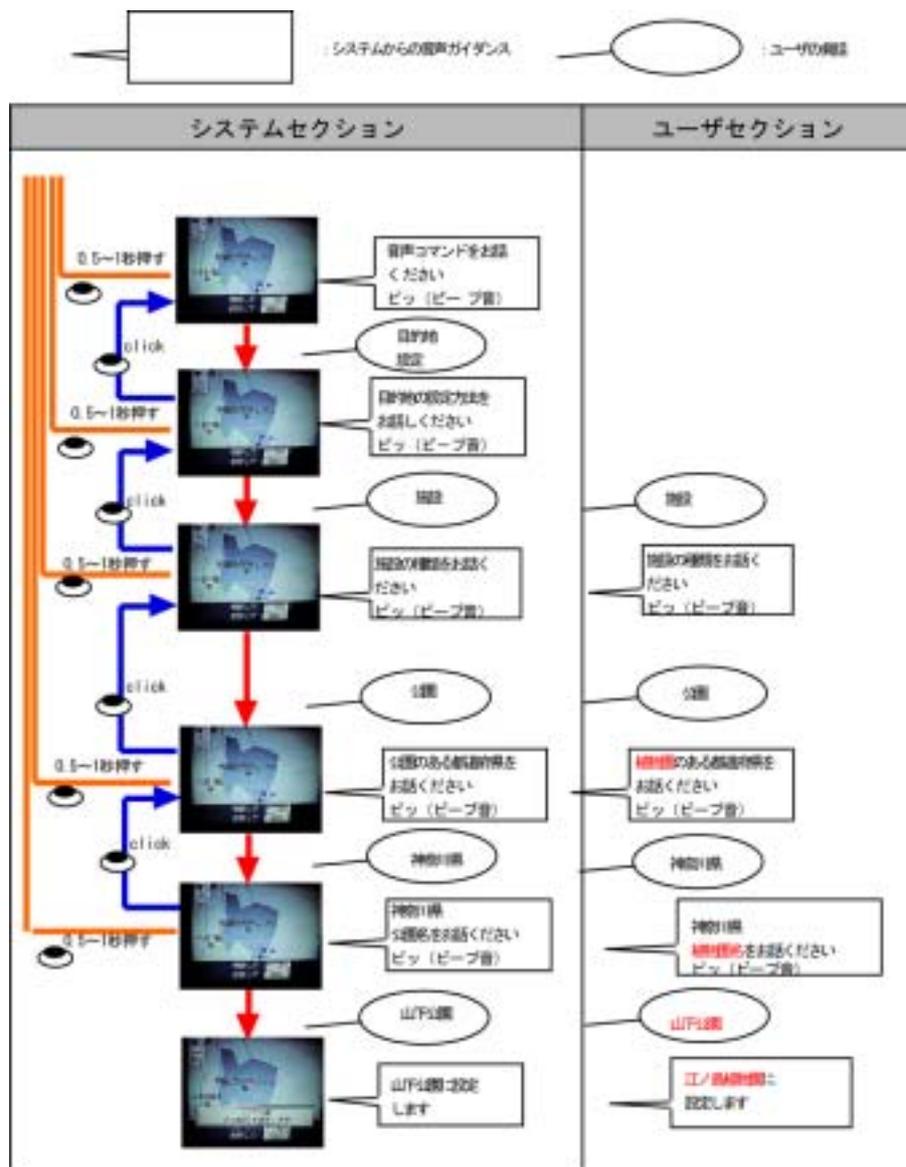


図 3.16 エラー・ループの例

### 3.6 考察

ストーリーボード評価システムを用いて音声認識型カーナビ操作時の様子を観察した結果、4種類のエラーが存在していた。これらのエラーが人間と機械間の円滑なコミュニケーションに負の影響を与えていると考えられる。これらのエラーのうち、特に不認識エラー、誤認識エラーの解決のためには認識率の向上が必要である。しかし、100%の認識率を発揮することは実際には困難であることを考えると、エラーが発生することを前提としたユーザインタフェースの設計が必須である。つまり、誤認識エラーのループを避ける工夫をする必要があると考えられる。

エラーのループを避けるためにはユーザが現在操作系を理解する必要があったと考えられる。D.A. ノーマンの「行為の7段階理論」[3-8]によれば、ユーザは評価の段階において「外界の状況の知覚」をし、「予期に基づいて知覚を解釈」し、「解釈を評価」する。今回のケースでは、誤認識エラーが操作系の分岐点において発生することによって、別の異なる操作系に移行したことをユーザが正しく理解できなかったために、エラーのループが発生していたことがわかった。これは、「外界の状況の知覚」が正しくなされていなかったと考えられる。システムは入力された物理情報をプログラムされたアルゴリズムに従って機械的処理を行うが、認識結果の正誤を決めるのはあくまで人間である。このことから、人間と同様な意味理解を機械に行わせる自動化も必要だが、一方ではエラーが発生した時点で、ユーザが状況を音声によって効果的に把握できるための工夫もユーザビリティの観点から重要であると考えられる。今回使用したカーナビの呈示音声には、直前の階層で入力した音声コマンドの入力結果がガイダンスの最初の部分にフィードバックされるようになっていた。しかし、この部分をきちんと聞いて理解できなかったことがループ発生の原因として考えられる。したがって、ループを避けるために、操作系の分岐点ではフィードバックの強化などの何らかの理解促進の方法を考える必要があるだろう。本章の結果を踏まえて、次章では改善策の提案を行う。

### 3.7 まとめ

本章では、逐次音声対話方式を用いて市販化されたカーナビを対象に、ユーザビリティという観点から、対話中の音声呈示においてどのような問題が起きているか検討を行った。まず、エラー率の計算とタスク開始から終了に至るまでの行動パターンを詳細に分析するために、階層構造をもつソフトウェア使用時の状態遷移を記述することができる評価システムとしてストーリーボード評価システムを構築した。次に構築したシステムを用いてユーザテストを行った。その結果、以下の知見を得た。

(1) 4種類のエラー（発話タイミングエラー、誤認識エラー、不認識エラー、誤操作エラー）が今回調査した市販カーナビ3機種全てにおいて発生していることがわかった。

(2) 操作系が切り替わる箇所で音声誤認識エラーが発生し、機器から呈示された復唱音声ユーザが理解できずに操作を進めた結果、音声誤認識エラーが連続的に発生してタスクが失敗に終わってしまうパターンが認められる。

(3)(2)より、直前の階層で入力した音声コマンドの入力結果がガイダンスの最初の部分にフィードバックされる部分をきちんと聞いて理解できたかどうか、タスク達成の成否に関わっていると考えられるので、誤認識エラーの連続発生を避けるために、操作系の分岐点においてはフィードバックの強化など理解促進の方法を考える必要がある。