

## マルチモーダル案内システム MMGS の実験評価

高橋 和子\* 竹澤 寿幸  
ATR 音声翻訳通信研究所 第四研究室  
(Phone) +81-774-95-1301  
(Fax) +81-774-95-1308  
(e-mail) { ktaka,takezawa }@itl.atr.co.jp

マルチモーダル案内システム MMGS の評価について述べる。このシステムは特に自然な対話を実現するために省略・照応の解決を効率よく扱える対話管理機構を持つことを特徴とする。被験者を用いて評価実験を行なった結果、例文を用いた対話について音声認識率 90%弱、平均文応答時間約 7 秒という結果を得た。また、課題付き自由対話についてはインタラクションが頻繁に起こる現象が確認され、MMGS の高いインタラクティブ性が例証された。本稿では、実験結果を示すとともに、マルチモーダル対話システムの評価方法についても論じる。

## Evaluation of Multimodal Guidance System (MMGS)

Kazuko TAKAHASHI and Toshiyuki TAKEZAWA  
ATR Interpreting Telecommunications Research Laboratories  
(Phone) +81-774-95-1301  
(Fax) +81-774-95-1308  
(e-mail) { ktaka, takezawa }@itl.atr.co.jp

We describe the evaluation of the multimodal guidance system (MMGS). This system has the dialogue mechanism of handling ellipsis/anaphora resolution effectively for the realization of natural dialogue. The experimental result of evaluation reports that the recognition rate is about 90% and response time is about 7 seconds in the dialogue using sample sentences. As for the free conversation for accomplishing some task, frequent interaction is observed, and it illustrates that MMGS is a dialogue system with high interaction. We also discuss the evaluation method for multimodal dialogue systems.

### 1 はじめに

1980年代後半から音声認識技術が飛躍的に進歩し、パッケージ化されたソフトウェアも販売されたことから音声認識技術がより身近なものになってきた。これらの中には、以前のようにユーザが単語を明確に区切りながら文法に適った文で堅苦しく対話をする必要はなく、かなり自然な話し方をしても受け付けられるようなものも多い。これを受けて、音声だけでなく画像入力や身振りでの入力などと組み

合わせてより使い勝手のよいインタフェースを備えたマルチモーダル対話システムの研究が注目をあびてきている [Tak94][KMMN94][NDI98][TM98].

マルチモーダル対話システムでは、一度言った語句の省略や指示代名詞を使った短い発話が多く見られ、インタラクションが頻繁に起こる。一方、目的指向型の対話システムでは省略された語句の補完をしなければシステムはデータベース検索が行なえない。この問題を解決する機能を持ち、しかも実時間で応答可能なシステムを構築するには音響モデルや言語モデルの設計だけでなく、システムの応答を生成する対話管理や問題解決機構の設計が大きく影響

\* 現在 三菱電機産業システム研究所  
(E-mail)takahashi@sys.crl.melco.co.jp

を及ぼす。

省略の補完に関してはこれまでも自然言語インタフェースを持つ多くの対話システムが取り組んでいるが [GAMP87][DSP91], これまで主として採用されている方式では、語句の意味の近さをかなり細かく定義する必要があり、データベースの語彙を増やそうと思うと非常に煩雑な改変作業が必要になる。

我々はこのような観点からマルチモーダルインタフェースを持つ観光案内システム MMGS (Multimodal Guidance System)[TM98] を研究開発してきている。このシステムは、音声および3次元グラフィックスで表示された地図画面上での指示によって入力されたユーザの質問に対して、システムはデータベースを検索し、写真やグラフィックス上の動作および音声によって応答するユーザ主導型の質疑応答システムである。特に、対話管理機構として発話履歴の新しい格納/検索方式を使って省略された語句を補完する機構を持ち、インタラクションを適切に扱っている。この方式では、過去の対話で言及された語句の中で、現発話に関するすべてのものを組み込んだデータとして現発話が示す対象が表され、発話ごとにその内容が更新される。我々は [TT99] でこの方式の形式化を行ない、この方式によれば実時間応答が可能であり、語彙の増加やタスクの変更にも比較的簡単に対応できることを示した。

一方、マルチモーダル対話システムの評価実験方法はまだ確立されておらず、特に対話機構の評価に関しては適切な指標がないのが現状である。目的指向型の対話システムの場合、ひとつの指標はタスク達成率であるが、音声認識率が悪いとタスクを達成できる可能性は極めて低く、対話機構の評価ができない。

本研究の目的はマルチモーダル対話システムの一つである MMGS の評価を行なうとともに評価方法、特に対話機構の評価方法を検討することである。

マルチモーダル対話システムの使いやすさの要因は大きく分けると操作性と応答性の二つと考えられる。操作性はいかに簡単に入出力の操作ができるかであり、音声認識率の高さや画像の見やすさ、扱いやすさなどである。応答性は応答の正確さ、速度、協調性などである。これらの評価は主観的なものも多く、ユーザの意見収集による判定にたよらざるをえない。そのために本研究でも、音声認識率、応答時間などの測定とともに被験者へのアンケートを実施した。また、ここでは応答性のひとつの指標とし

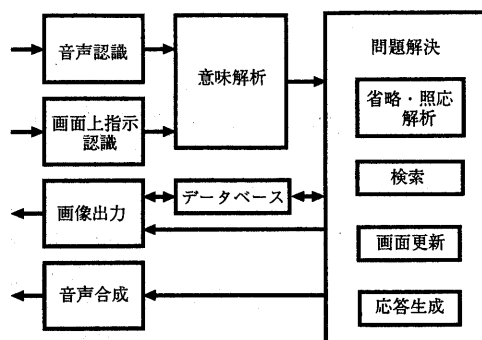


図 1: MMGS のシステム構成

「インタラクティブ性」を導入し、発話がどのくらい連続して認識されているか、省略・照応がどのくらい頻繁に出現するかを測定し、これによってインタラクションがうまくいっているかどうかを判断した。

その結果、例文を使った対話での音声認識率は90%弱、平均応答速度が約7秒であり、アンケートでも良い評価を得た。課題付き自由対話については操作性、応答性とも問題点が多かったが、インタラクティブ性は高かった。

本稿は以下のように構成される。まず、第2章で MMGS の構成を示す。次に、第3章で MMGS の対話管理機構について述べ、第4章で評価実験について述べる。最後に第5章で結論を述べる。

## 2 MMGS の構成

MMGS の構成を図 1 に示す。

ユーザの入力は音声および画面上のポインティングを認識する認識部を通った後、意味解析部によって意味解析の行われた意味表現が生成される。音声認識手法には HMM-LR, 言語モデルには CFG 規則および前終端記号の bigram を使っている。問題解決部ではこの意味表現をもとに省略や照応の解決を行なってデータベースの検索を行い、画面の更新情報や応答文を生成する。それを音声合成して出力したり3次元グラフィックスや写真によって表示し、必要に応じ画面の回転や点滅などを行なう。また、応答が返るまでの待ち時間を楽しませるためにキャラクターアニメーションを画面上に出す(図2)。データベースは階層型オブジェクト構造を持っており、各オブジェクトは属性として意味情報と空間・形状情

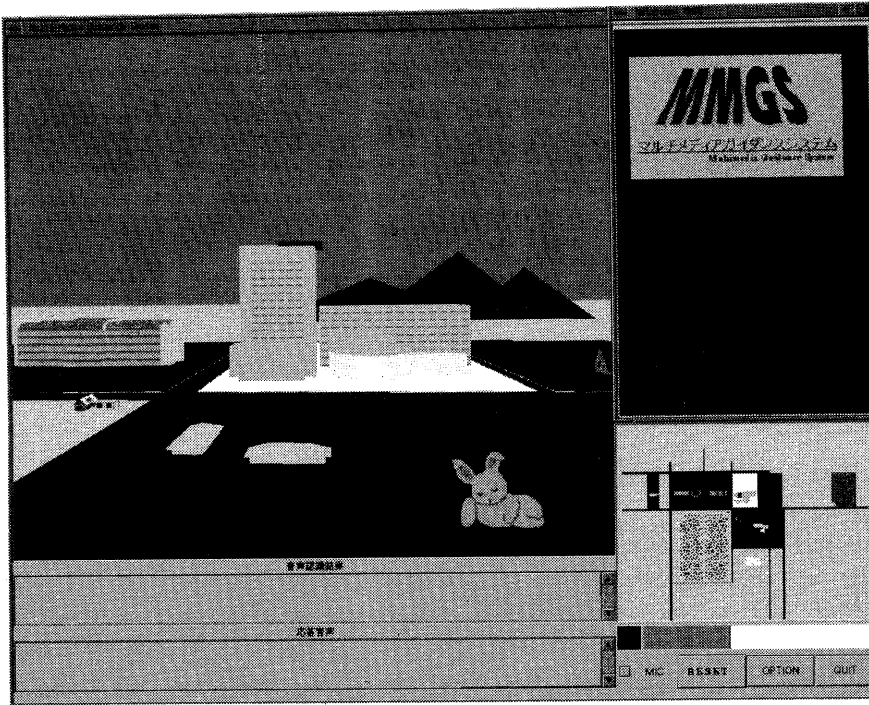


図 2: MMGS の表示例

報を合わせて持つ。

### 3 対話管理機構

MMGS の対話管理機構について簡単に述べる。  
(詳細は [TT99] を参照。)

#### 3.1 対話例

簡単のため各発話は単文 1 つで話者交代が生じるものと仮定する。

以下に対話例を示す。

##### 例 1. けいはんな地区の案内

U1: これはなんですか [画面上を○で囲む].

S1: 「けいはんなプラザ」です。

U2: 喫茶店がありますか。

S2: 「ライン」があります [写真を見せる].

U3: ATR にはありますか。

S3: カフェテリアがあります [写真を見せる].

[ ] 内は画面上の操作を示す。U はユーザの発話、S はシステムの発話をそれぞれ表す。U2 では「けいはんなプラザに」が、U3 では「喫茶店は」がそれぞれ省略されている。

一般に質疑応答型システムでは、ユーザが質問しそれに続いてシステムが答えるという隣接ペアが存在するものとしてよい。各隣接ペアごとに Q 話題および A 話題を対応させる。これらは現発話で直接言及された対象だけでなく、それが参照する過去の発話で言及されたものも含む。

#### 3.2 データ型

問題解決部への入力は (*class, action, object*) の 3 項組で表現される。*class* は質問の種類を、*action* は主動作の種類をそれぞれ表す。*object* はその文で言及されている対象物を表し、名前 (*label*) および素性 (*ptr*) の  $n$  個の ( $n=0$  を含む) 有限列であり、その構造は *class, action, label* によって決まる。各素性は素性名 (*fn*) と素性値 (*fv*) から成る。素性名は意味的な種類を表す。たとえば、喫茶店 (*coffee-shop*) という *label* を持つ *object* に対しては場所 (*loct*) という素性名を持つ構造がつけられる。主語や目的語が省略された発話では素性名ができて素性値が決まらない (すなわち変数のままである) 構造ができる。以後、値の決まっていない部分は '?' で始める文字列

で表す。

### 3.3 Q 話題の生成

入力である意味表現 (*Rep*) が変数を含む場合、直前の隣接ペアで生成された A 話題 (*Ainfo*) を使って補完を行い、現在のユーザ質問に対応する Q 話題 (*Qinfo*) を生成する。object のある素性は判明しているが label が不明な場合、*Ainfo* に素性名の一致するものがあれば、その素性値といれかえる。object 自体が変数の場合、*Ainfo* そのものを代入する。いずれでもない場合は、データベースを検索して同じ構造を持つオブジェクトを探し出し、補完する。

例 2. 例 1 の U3 「ATR にはありますか」

```
Rep = (yn-q, exist,  
      (?LABEL, (loct, ATR), ?FEATURE))
```

直前の発話で「けいはんなプラザの喫茶店」が言及されているので

```
Ainfo = (coffee-shop, (loct, keihanna-plaza))
```

が成り立っており、ここから rhein(ライン) へポインタがはられている。

この時、object の変数部分 ?LABEL および ?FEATURE を補完して *Qinfo* を生成する。ここで、*Ainfo* には loct という素性名があるので、この素性値を置き換えて、

```
Qinfo = (coffee-shop, (loct, ATR))
```

を得る。同時に ?LABEL には coffee-shop が代入され、?FEATURE には NIL(情報なし) が代入される。

例 3. 例 1 の U3 の代わりに U3' 「広いですか」という質問がきた場合

```
Rep = (yn-q, wide, ?OBJECT)
```

この場合、?OBJECT の構造には制約がないので、*Ainfo* そのものを代入してよいが、*Ainfo* からポインタがはられているので、その先の rhein を代入し、*Qinfo* = rhein を得る。

### 3.4 A 話題の生成

ユーザの質問に対応した Q 話題が生成されると、システムはこれに追加、訂正などの修正を加えた A 話題を生成し、これを適当な形にして回答とする。ユーザの質問の種類によって A 話題の生成方法も変わるが、ここでは yn-question によるオブジェクト

の存在の質問の場合を例として *Qinfo* から *Ainfo* を生成する方法を述べる。

対応するオブジェクトはデータベースに存在する場合は、*Qinfo* に対する具体値である instance を返すので、*Qinfo* から instance へのポインタをはったものを *Ainfo* とする。ない場合は、NIL とする。

例 4. 例 1 の S3 「カフェテリアがあります」

```
Qinfo = (coffee-shop, (loct, ATR))
```

データベースを検索すると対応するオブジェクトとして cafeteria が見つかるのでここにポインタをはる。

### 3.5 特徴

*Qinfo/Ainfo* の生成の際、唯一履歴として参照にするのが *Ainfo/Qinfo* だが、この参照では語句の長さや構造にかかわらずデータへの一度のアクセスによって補完すべき語句を見つけることができる。また、これまでの発話すべての履歴を持つ必要がないので探索も速い。さらに、音声認識がうまく行われなかった場合でも意味表現として、class, action さえ認識できていれば、不明な部分の補完が可能である。なお、以上では省略について述べたが、照応に関しても同様に扱える。

## 4 評価実験

被験者を使って MMGS の評価実験を行なった。

被験者は 19 人(男 11 人, 女 8 人)である。年代は 10 代~40 代, 職業は学生, 大学教員, 研究開発従事者であり, 全員コンピュータの使用経験を持っていた。

### 4.1 実験方法

システム概要と使用方法を説明した後、被験者に例文による対話および課題付自由対話を行なってもらい、ログファイルをとった。話者適応やシステムの使い方の練習、音声入力の練習はいずれも行っていない。また、被験者には実験終了後アンケートを記入してもらった。

例文による対話では、2 場面 15 文の発話を行なってもらった(約 10 分)。自由対話では、課題を一つ指定し、10 分程度の時間内で課題に即した対話をしてもらった。原則として言い回しの制限はしないが、あまりに認識が悪い場合は適宜誘導することにした。

	発話数	正解 認識数	平均文 認識率 (%)	平均応答 時間 (秒)
例文対話	330	289	87.6	6.9
自由対話	389	200	41.4	7.5

表 1: 音声認識率

	(秒)		
処理	全体	認識成功時	失敗
音声認識	3.1	3.0	3.6
意味解析	3.4	-	-
問題解決	0.4	-	-
合計	6.9	6.8	7.5

表 2: 例文対話の平均応答時間

## 4.2 動作環境

MMGS は SGI Octane および Linux/PC Endeavor Pro300 上に実装されている。前者には音声画像処理部分、後者にはその他の処理部分が実装され、両者がネットワークを介して通信している。音声認識部と意味解析部はサイズは約 300 語の同一辞書を持つ。データベースのオブジェクト数は約 50 個である。

## 4.3 結果および考察

### 4.3.1 システムログ結果および考察

ユーザには 1 発話ごとにマイクを切断するように指示し、そのタイミングを発話終了時刻とした。発話終了時刻からシステムの応答開始時刻までを応答時間とし、切り忘れやあまりにタイミングの遅いものは応答時間を出す集計から除外している。

例文対話については音声認識率、応答時間、繰り返し質問の回数を測定した。自由対話については音声認識率、応答時間、画面表示の操作回数に加えインタラクティブ性を測定した。これらの結果に性差、年齢差はほとんど見られなかった。

### 音声認識率および応答時間

表 1 に音声認識率および応答時間を、表 2 に例文対話の応答処理における各ステップの実行時間をそれぞれ示す。

例文対話では平均文認識率は 87.6%、自由対話では 41.4% であった。自由対話で認識率が低いのは

失敗原因	件数	割合 (%)
認識誤り・文法未対応	129	62.0
文型未対応・データ不足	58	27.9
問題解決アルゴリズムの不備	16	7.7
解析結果書き換え	5	2.4
合計	208	100.0

表 3: 応答処理失敗の原因

MMGS の語彙数が基本的に少ないためであり、今後整備していく予定である。

平均応答時間は例文対話、自由対話でそれぞれ 6.9 秒、7.5 秒であり、ユーザには少々遅いという印象を与えている。この主原因は、意味解析処理において認識結果をテキストデータとして SGI マシンからネットワークを介して PC に送っていることと、汎用の意味解析プログラムの出力を MMGS データベースの問合せ用のものに書き換えているためである。この部分の本来の処理は 0.1 秒もかからないと見積もられ、意味解析プログラムを MMGS 専用のものに修正し同一マシン上で実装することによって 1 桁以上の高速化が期待される。また、音声認識処理では、現在 150 万幅で行なっている認識のビーム幅を減少すれば処理速度は半減可能なのかわかっている。問題解決処理では、相当な数の場合分けをして場合ごとに応答を生成しており、改良してもオーダをさげるほどの大きな効率アップはのぞめないが、場合分けの順序を工夫したり応答パターン数を増やすことで、ある程度的高速化は可能である。以上によって、全体の応答時間は少なくとも 3 分の 1 程度に削減可能と考えられる。

### 応答処理失敗の原因

例文対話では繰り返し質問数は 15 発話中平均 2.1 回、最多 7 回、最少 0 回であった。繰り返しの原因はすべて音声認識の失敗である。

自由対話における応答処理失敗の原因を表 3 に示す。

この表で、認識誤り・文法未対応は音声認識部の失敗によるものであり、認識できない文型、認識誤り、ノイズ、ポーズのために複数の文と認識されたことなどが含まれる。

文型未対応・データ不足には、データベースにデータが無い、問題解決部で対応できない文型が入力されたことなどが含まれる。

問題解決アルゴリズムの不備では、認識の失敗によって Q 話題/A 話題の履歴が変わり、正しい情報を与えることができなかつたことによるものが多かった。音声対話では認識失敗は当然起こるものであり、今後は認識失敗に対しても対応できるような履歴管理が行なえるようにアルゴリズムを改良していく必要がある。

解析結果の書き換えには、ある種の質問に対する処理が未対応であったことなどが含まれる。上で述べたように意味解析部を MMGS 専用のものにカスタマイズすればこの書き換え処理は不要になる。

#### 画面上表示操作

MMGS には話題となっている対象物が見えやすい位置に画面が移動する機能があり、マウス操作によってユーザ自ら視点を動かすこともできるのだが、自由対話において実際に画像を動かした者は 1 名、ポインティングをした者も 3 名のみだった。これは、どの被験者も音声入力に手一杯で、画像入力まではできなかったためと考えられる。課題に画像入力を特に必要とするものがなかつたのも原因の一つだが、「マルチモーダル入力」を使いこなすためには、音声認識率の向上および問題解決部の補強が必須だろう。

#### インタラクティブ性

我々はインタラクティブ性の指標のひとつとして、以下のものを考えた。

- 正しく認識された発話の中で連続出現する数  
- 多いほどインタラクティブである
- 省略・照応の数  
- 多いほどインタラクティブである

自由対話について表 4 に各被験者の 1 セッションにかかった時間、発話数、インタラクティブ性などの測定結果を示す<sup>1</sup>。各項目は以下の通りである。

1. 1 セッションの発話数
2. 正しく認識された発話数
- 2a. このうち連続出現するものの数の列
- 2b. このうち省略・照応のある発話数
3. セッションの時間 (システムダウンの時間は除外)
4. 文認識率 (= 項目 2 / 項目 1)
5. 省略・照応の出現頻度 (= 項目 2b / 項目 2)

<sup>1</sup>被験者 A はログ不良のため無視した。

インタラクティブ性に関しては項目 2a,2b が相当する。項目 2a で、たとえば被験者 C では、2 個連続してうまく応答された発話列と 3 個連続してうまく応答された発話列があったことを表す。この項目に関しては、2,3 の質疑応答が続くとその後応答処理に失敗してしまうというパターンが多く、あまりインタラクションがうまくいったとは思われない。しかし、項目 2b が示すように、ほとんどの被験者が (おそらく無意識のうちに) 省略・照応を使った発話をしており、これらの解決を当然あるべき機能と考えていることがわかる。このことから、MMGS の対話管理機構が適当なものであり、MMGS が高いインタラクティブ性を持つことが例証された。

#### 4.3.2 アンケート集計結果および考察

アンケートの集計結果を以下に示す。

##### 音声入出力

例文対話に関してはおおむね好印象であったが、自由対話についてはほとんどの人が表現したいことが伝えられずストレスを感じたようである。

##### 画像入出力

動きの速さ、大きさなどについては比較対象がないため答えにくかったようで、特に意見は見られなかった。

##### 対話機構

自由対話では全員がタスク達成できず、認識で失敗していたため話しにくいという印象が濃い。応答文の表現については特に不自然な感じはないという評価を得た。

##### 全体

応答時間は若干長めという意見が多かった。これについては 4.3.1 節で述べたような改良によって高速化可能であり、今後の課題である。また、うさぎのアニメーションキャラクタは好評でありバリエーションがほしいという意見も多かった。

さらに、マルチモーダル対話システムの将来性についてだが、19 名中 15 名が案内システムとして適当な形態だと回答している (残りの 4 名はどちらともいえないと回答)。また、19 名中 17 名が使いたいと回答し、そのための条件として音声認識率の向上を条件にあげた者が最も多かった (14 名)。使いたくないという意見の中には、機械に向かって話しかけることへの抵抗というものがあった。今後、普段計算機に接することのないユーザを含む幅広い層

被験者	1	2	2a	2b	3	4 (%)	5 (%)
B	15	8	4	5	373	53	63
C	17	8	2,3	2	304	47	25
D	19	6	2	1	412	32	17
E	9	6	3	1	218	67	17
F	19	7	5	4	525	37	57
G	37	18	6,3,2,2,2	5	729	49	28
H	28	23	2,2,4,14	12	737	82	52
I	35	11	2,3,2	2	686	31	18
J	31	11	4	5	744	35	45
K	38	14	0	7	917	37	50
L	23	14	2,3	3	578	61	21
M	14	7	3	0	324	50	0
N	32	14	2,2,2	3	825	44	21
O	26	4	0	0	594	15	0
P	47	19	2,2,3	7	832	40	37
Q	32	8	3,2	3	708	25	38
R	26	11	2,3,3	5	565	42	45
S	35	11	2	4	848	31	36
平均	26.8	11.1	-	3.8	606	41.4	34.2

表 4: 自由対話の結果

にマルチモーダルインタフェースを標準的なものとして浸透させるためには、機械に向かって話しかけるイメージを改善する必要がある。しかし、そのために、できるだけ人間に近いものをつくるのがよいのか、それとももっと別の形で実現すべきなのかは今後のユーザインタフェースの研究の課題である。

#### 4.4 対話システムの評価と実験方法

対話システムの評価、特にインタラクティブ性の評価に関しては、基準となる指標が今のところ存在しない。

目的指向型対話では、タスク達成率がひとつの指標となるが、対話システムが(分野は限られても)十分実用に耐えるほど鍛えられていない場合は、初心者タスク達成率は極端に低くなってしまふ。[NDI98]では、話者適応とともに練習セッションをもうけて被験者がある程度システムを使いこなせるようになってから実験を開始している。このような方法だと被験者のストレスも少ないし良いデータが得られるが、一方、使い方を限定してしまい、型にはまった言い回しが多くなって自然な対話の評価を行なうという

今回のような実験の目的に反する。

また、マルチモーダル対話システムから対話管理部分のみを切り出すのは難しいことから、音声認識を切り離しマウスによるメニュー選択やキーボードからのテキスト入力によって対話を行ない、問題解決部分のみの評価データを測定することも考えられる。しかし、この場合入力が音声・画像インタフェースを通した時のものと同じものになるかという点では疑問が残る。

また、使い勝手に関しては複数のシステムを比較しないで評価を行なうのは非常に困難であり、音声入力のみ、あるいはメニュー形式のインタフェースを持つシステムとの比較も有効だろう。Litmanらは、対話システム全体の性能評価として、話者交代の回数、質問の種類、アンケートによるユーザの満足度などをパラメータとして計算する方法を使い、[LPW98]で応答戦略の異なる二つのシステムについて評価実験を行なった。しかし、彼らが実験に用いたのはいずれもシステム主導型の対話システムであり、MMGSのようにユーザが主体的に発話を行なうような場合にはそのままでは当てはまらない。

マルチモーダル対話システムの評価実験方法については我々が今回とった方法が最良とはいえ、今後も検討が必要である。

## 5 おわりに

マルチモーダル案内システム MMGS の評価について述べた。このシステムは特に省略・照応の解決を主体とする対話管理機構を持つことを特徴とし、インタラクティブ性の高い自然な対話の実現を目標としている。実験の結果、例文を使った対話については音声認識率、応答時間を含めユーザにはほぼ満足のいくものになった。課題付き自由対話についてはまだまだ問題点が多いが、インタラクションが頻繁に起こる現象が確認され、MMGS が高いインタラクティブ性を持つシステムであることが例証された。

今後は実験結果をシステムにフィードバックするとともに、評価実験方法特にインタラクティブ性の評価方法について検討していく予定である。

## 謝辞

MMGS の実装に関しては田中吾郎氏、塩崎純氏、馬場武雄氏に御協力いただきました。ここに記して感謝します。

## 参考文献

- [DSP91] M. Dalrymple, S. Shieber and F. Pereira. *Ellipsis and Higher-Order Unification*, *Linguistics and Philosophy* Vol.14, No.4, 1991.
- [GAMP87] B. J. Grosz, D. Appelt, P. Martin and F. Pereira. *TEAM: An Experiment in the Design of Transportable Natural Language Interfaces*, *Artificial Intelligence*, vol.32, no.2, pp.173-244, 1987.
- [KMMN94] 神尾広幸, 松浦博, 正井康之, 新田恒雄. マルチモーダル対話システム *MultiksDialog*. 電子情報通信学会論文誌, vol.J77-D-II, no.8, pp.1429-1437, August, 1994.
- [LPW98] D. J. Litman and S. Pan and M. A. Walker. *Evaluating Response Strategies in a*
- Web-Based Spoken Dialogue Agent, In *COLING-ACL 98*, pp.780-786, 1998.
- [NDI98] 中川 聖一, 傅田 明弘, 伊藤 敏彦. マルチモーダル観光案内対話システム. 人工知能学会誌, vol.13, no.2, pp.241-251, March, 1998.
- [Tak94] 竹林 洋一. 音声自由対話システム TOSBURG II - ユーザ中心のマルチモーダルインタフェースの実現に向けて -. 電子情報通信学会論文誌, vol.J77-D-II, no.8, pp.1417-1428, August, 1994.
- [TM98] T. Takezawa and T. Morimoto. *A Multimodal-Input Multimedia-Output Guidance System: MMGS*, In *Fifth International Conference on Spoken Language Processing*, vol.2, pp.285-288, 1998.
- [TT99] 高橋 和子, 竹澤 寿幸. マルチモーダル対話システムにおけるインタラクション機構. *インタラクション'99*, pp.89-96, 情報処理学会, 1999.

## 付録. 評価実験で使用した対話の例

### 例文対話 (一部)

1. 奈良公園の観光コースを教えてください。
2. 時間はどのくらいかかりますか。
3. 大仏は何時から見られますか。
4. いくらかかりますか。
5. 近くで昼食は食べられますか。
6. 奈良公園の中にはないんですか。
7. 大仏殿から歩いて行けますか。
8. これは何ですか [茶店の隣の建物を○で囲む]。
9. 奈良のみやげものを教えてください。
10. どうもありがとうございます。

### 自由対話の課題 (例)

あなたは今近鉄奈良駅におり、これから奈良県新公会堂で行なわれる会議に出席するためそちらに行かねばなりません。あなたは行き方についてはまったく知りません。会議が早く終われば奈良公園を歩いて近鉄奈良駅まで帰ろうと思っていますが、遅くなればバスを使おうと思っています。