

# 車載用音声対話システムにおける ユーザ負荷を考慮した対話戦略の検討

山岡 将綺<sup>1,a)</sup> 原 直<sup>1,b)</sup> 阿部 匡伸<sup>1,c)</sup>

**概要:** 運転中の利用に適した対話方法を提案するため、ユーザ負荷の観点で対話戦略を評価した。対話の主導権の方式と確認発話の方式を組み合わせ、客観評価実験である計算機シミュレーションを行った。シミュレーションの条件として、システムがユーザに余裕があるときのみ発話する条件と、ユーザが正しく聞き取れる割合の条件を設定した。また、これらの方式を実装した音声対話システムを利用した主観評価実験を行った。これらの評価から、ユーザ主導の対話戦略は認識率が高いとターン数やタスク達成時間が少なく優れているという結果を得た。一方、システム主導や混合主導の対話戦略は認識率が低い場合にターン数やタスク達成時間がユーザ主導の対話戦略よりも低く優れているという結果を得た。余裕があるときのみ発話する場合は、タスク達成までに時間がかかるため効率が悪いが、ユーザが感じる負担はほとんどないという結果を得た。

**キーワード:** 対話戦略, 対話の主導権, 確認発話, 音声対話システム

## Influence analysis on user's workload in a spoken dialog strategy for a car navigation system

MASAKI YAMAOKA<sup>1,a)</sup> SUNAO HARA<sup>1,b)</sup> MASANOBU ABE<sup>1,c)</sup>

**Abstract:** We assess dialog strategies from user's workload standpoint to suggest usage of spoken dialog systems while driving. We evaluate several dialog strategies, which are combinations of methods of dialog initiative and confirmation manner, with objective evaluation method using computer simulation. Two conditions of the simulation are set up; One is that the system speaks if the user has leeway to talk with the system, and the another one is that the system speaks even if the user will be fail to recognize the system's. We also evaluate spoken dialog systems applying these methods with subjective evaluation method. As a result of the evaluations, user initiative strategy has advantages in lower turn number and lower task completion rate than both system initiative strategy and mixed initiative strategies when cognitive rate is high. The result also shows that system initiative strategy and mixed initiative strategy have advantages in lower turn number and lower task completion rate than user initiative strategy when cognitive rate is low. Additionally, the result shows that the method, which system speaks only when user has enough time driving operation, makes user's workload low, however, it need more time to complete tasks.

**Keywords:** dialog strategy, dialog initiative, confirmation speech, spoken dialog system

### 1. はじめに

近年、音声認識技術や音声合成技術が成熟してきたこと

により、これらの技術を応用して、人間と機械との間でコミュニケーションをとる音声対話システムが日常生活で見られるようになった。例えば、スマートフォンでのインターネット音声検索や、カーナビゲーションシステムでの目的地設定などに使用されている。音声を用いて機械の操作をすることの利点として、画面を注視する必要がないこと（アイズフリー）や、入力の際にキー操作が必要でない

<sup>1</sup> 岡山大学大学院自然科学研究科  
Graduate school of Natural Science and Technology,  
Okayama University

a) yamaoka@a.cs.okayama-u.ac.jp

b) hara@cs.okayama-u.ac.jp

c) abe@cs.okayama-u.ac.jp

こと（ハンズフリー）があげられる [1]。特に、他の作業をしながらでも入力できるため、運転中でも脇見運転をせず、安全に使用することができると期待されてきた。

しかし、音声対話システムはユーザに認知負荷がかかり、運転に関わる反応が遅れてしまう恐れがある。その原因として考えられるのは、音声認識器の誤認識である。人間とコンピュータが音声で対話をする場合、音声認識器の誤認識は避けられない。また、インタフェースの適切な設計法は十分には研究されておらず、システムが誤認識した際の回復に多くの課題が残っている [2]。誤認識したまま対話を進めることを防ぐために、ユーザの入力ごとに確認発話を行うシステムが多いが、確認発話のために、タスク達成までのターン数や時間が増えてしまう。ユーザが煩わしく感じてしまうという問題だけでなく、運転中に利用するシステムであれば、運転者の注意が運転行為からシステム操作に向かってしまい、交通事故の要因となる恐れもある。以上のことから、ユーザにかかる負荷をできるだけ小さくした音声対話システムが必要と考える。

本報告では、運転中のユーザへの対話負荷が少ない戦略を明らかにすることを目的とする。まず、運転中に適した対話戦略を明らかにするために、対話の主導権の方式と確認発話の方式を組み合わせ、シミュレーション実験を行う。シミュレーション実験は、運転中の利用者に余裕がある場合とない場合を想定し、システムがユーザに余裕がある場合のみ発話する余裕時発話条件と、ユーザの肯定文と否定文の発話正答率を 80%にするユーザ誤認識条件の 2 つの条件で行う。評価項目は、ターン数平均、タスク達成時間平均とタスク達成率である。次に、これらの方式を実装した音声対話システムを利用した主観評価実験を行う。評価項目は、効率性、非負担度と満足度である。

本稿の構成は以下の通りである。まず、2 節では本研究で用いた対話システムの概要について述べる。3 節では対話システムを用いた評価実験について述べ、4 節では主観評価実験について述べ、5 節で結論を述べる。

## 2. ユーザ負荷を考慮した対話システムの構築

### 2.1 タスク達成型の対話システム

本報告では、タスク達成型の対話システムを使用する。タスク達成型の対話システムとは、ユーザが持つ目的を達成するために、対話を行うシステムである。ユーザはいくつかの情報をシステムに入力し、その情報をもとに対話を進めながらタスク達成を目指す。本報告での音声対話タスクは、カーナビゲーションシステムで用いられる目的地設定とする。ユーザは県名、市名、目標物の 3 つのスロットを対話によって埋めていき、全てのスロットが埋まればシステムは目的地設定をして対話は終了とする。

### 2.2 安全運転に配慮した対話戦略

一般的な音声対話システムでは対話時間を短くするために、ユーザが音声を入力するとすぐに応答を返すシステム

が多い。しかし、ユーザは車の運転時には運転に集中しており、対話システムとの対話を望んでいない場合が考えられる。また、ユーザがシステムの応答を聞き間違えることも考えられる。そこで、本報告では運転者に余裕がある場合のみシステムが発話する方式を提案する。余裕がある場合には自動車が一時停止しているときや一定速度で走行中のときが考えられ、余裕がない場合には加減速中やハンドル操作をしているときが考えられる。これらの運転状態はカーナビゲーションシステムやスマートフォンに搭載されている加速度センサー、GPS により観測可能である。

提案する対話戦略では、余裕がない場合にはシステムからの応答が停止されるため、ユーザが対話に邪魔されることなく運転に集中することができる。余裕がある場合ならばユーザはシステムとの対話に集中することができるため、ユーザがシステムの応答を聞き間違えることが少なくなると期待できる。これに加えて、ユーザは落ち着いて話すことができるため、システムで認識しやすい音声をユーザが発声することも期待される。

本報告では、ユーザが入力する内容を一つ一つ指定する戦略をシステム主導、ユーザが一度の発話で内容を入力する戦略をユーザ主導、最初の発話はユーザが自由に入力して 2 回目の発話からはシステム主導とする戦略を混合主導と呼ぶ。これら 3 つのシステムの制御方法の対話例を図 1 に示す。また、ユーザが入力した内容を単純に聞き返す方法を明示的確認発話と呼ぶ。それに対して、ユーザが入力した内容を含めつつ次に入力してほしい内容の質問をする方法を暗黙的確認発話と呼ぶ。実装した対話戦略を以下に示す。

#### (1) システム主導

- (a) 明示的確認発話
- (b) 暗黙的確認発話

#### (2) ユーザ主導

#### (3) 混合主導

- (a) 明示的確認発話
- (b) 暗黙的確認発話

## 3. ユーザ発話シミュレータを用いた対話システムの客観的評価実験

### 3.1 ユーザ発話シミュレータ

本報告では、ユーザ負荷を考慮した対話システムをシミュレーションするために、次の 2 つの条件を設定する。1 つ目は、システムがユーザに余裕がある場合のみ発話する条件である。以下、余裕時発話条件と呼ぶ。この時ユーザの余裕状態（余裕あり、又は、余裕なし）は等確率で変化し、システムによるユーザの状態判定は 5 秒毎に行う。2 つ目は、運転中という状態ではユーザが合成音を正しく聞き取れない可能性があると考え、ユーザの肯定文と否定文の正答率を 80%にする条件である。以下、ユーザ誤認識条件と呼ぶ。

```
システム主導（明示的確認発話）
sys： 県名を教えてください
usr： 岡山県
sys： 岡山県でよろしいですか
usr： はい
sys： 市名を教えてください
usr： 玉野市
sys： 高梁市でよろしいですか
usr： いいえ，玉野市
sys： 玉野市でよろしいですか
usr： はい
sys： 目標物を教えてください
usr： コンビニ
sys： コンビニでよろしいですか
usr： はい
sys： 岡山県玉野市のコンビニに設定します
ユーザ主導
sys： どこに行きたいですか
usr： 岡山県玉野市のコンビニ
sys： 岡山県高梁市のコンビニでよろしいですか
usr： いいえ，岡山県玉野市のコンビニ
sys： 岡山県玉野市のコンビニでよろしいですか
usr： はい
sys： 岡山県玉野市のコンビニに設定します
混合主導（暗黙的確認発話）
sys： どこに行きたいですか
usr： 岡山県玉野市
sys： 岡山県高梁市の目標物を教えてください
usr： いいえ，岡山県玉野市
sys： 岡山県玉野市の目標物を教えてください
usr： コンビニ
sys： コンビニでよろしいですか
usr： はい
sys： 岡山県玉野市のコンビニに設定します
```

図 1 システムの制御方法  
Fig. 1 Control method of system.

### 3.2 実験条件

2.2 節の 5 つの対話戦略を，余裕時発話条件，ユーザ誤認識条件の 2 つの条件でそれぞれ 1000 回のシミュレーションを行った。タスク達成に必要なユーザ発話は県名，市名，目標物の 3 つである。音声認識部でのシステムの想定単語認識率は 60% から 100% まで 5% 刻みとする。ユーザやシステムの話す速度は 8 モーラ/秒とし，発話と発話の間隔は 1 秒として計算する。

### 3.3 評価尺度

本実験では以下の 3 つの評価尺度に基づく評価を行った [3]。

#### (1) ターン数平均

ユーザがタスクを達成するまでにユーザとシステムが発話した回数を計測した尺度である。30 ターンまでにタスクが達成されない場合はタスク未達成として対話を打ち切る。一般的に，ターン数は少ないほど効率的でよい対話戦略と言える。

#### (2) タスク達成時間平均

タスク達成までにかかる時間を計測した尺度である。

余裕時発話条件では 100 秒，ユーザ誤認識条件では 60 秒までにタスクが達成されない場合はタスク未達成として対話を打ち切る。タスク達成時間が短いほどよい対話戦略と言える。

#### (3) タスク達成率

タスクを一定回数実施して，実際にタスクが達成された割合を示す尺度である。本報告では，ターン数とタスク達成時間のいずれかで打ち切られた場合を，タスク未達成と判定して対話を打ち切る。打ち切られた対話は，ターン数やタスク達成時間平均の計算には利用していない。

### 3.4 実験結果

余裕時発話条件のシステムの認識率に対するターン数平均とタスク達成率の関係を図 2 に，タスク達成時間とタスク達成率の関係を図 3 に示す。図 2 を見ると，ターン数平均はユーザ主導，混合主導，システム主導の順に増えていく。また，暗黙的確認発話より明示的確認発話の方がターン数が多くかかる傾向にある。システムの認識率に対するターン数平均の傾きは対話戦略によらずほぼ一定であり，システムの認識率の変化による順序の交替も起こっていない。

図 3 を見ると，暗黙的確認発話を用いたシステム主導と明示的確認発話を用いた混合主導のタスク達成時間平均は同程度となっている。その他の対話戦略の順序はターン数平均と同じであることが分かる。システムの認識率が下がっても，システム主導や混合主導のタスク達成率が下がることはないが，ユーザ主導はシステムの認識率が 75% 以下になるとタスク達成率が大幅に下がる。

ターン数平均やタスク達成時間平均の傾きが一定でありながら，タスク達成率が大幅に下がる理由として，システムの認識率が下がると対話のターン数や時間が長くなり，打ち切り条件に該当する対話が大幅に増えたためだと考えられる。以上より，ユーザ主導はシステムの認識率が 75% 以上のときはターン数やタスク達成時間が少なくてすむが，システムの認識率が 75% 以下になると大幅にターン数やタスク達成時間が増え，タスク達成回数が少なくなってしまうことが分かる。一方で，システム主導や混合主導はユーザ主導よりもターン数やタスク達成時間はわずかに増えてしまうが，システムの認識率が下がってもターン数やタスク達成時間が大幅に増えることはなく，タスクはほぼ達成できる。

次に，ユーザ誤認識条件のシステムの認識率に対するターン数平均とタスク達成率の関係を図 4 に，認識率に対するタスク達成時間平均とタスク達成率の関係を図 5 に示す。図 4 を見ると，ターン数平均の順序は余裕時発話条件と同じであることが分かる。図 5 を見ると，タスク達成時間の順序はターン数平均と同じである。タスク達成率はシステムの認識率の低下に伴って低下してしまう。このことから，タスク達成率を一定以上にするためには，システム

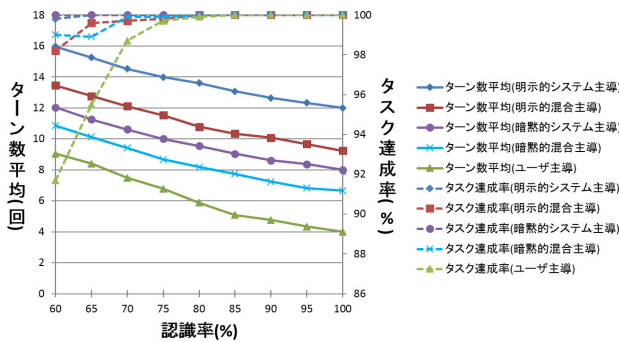


図 2 ターン数と達成率 (余裕時発話条件)

Fig. 2 Turn and completion rate of condition systems speak users only when users afford speak.

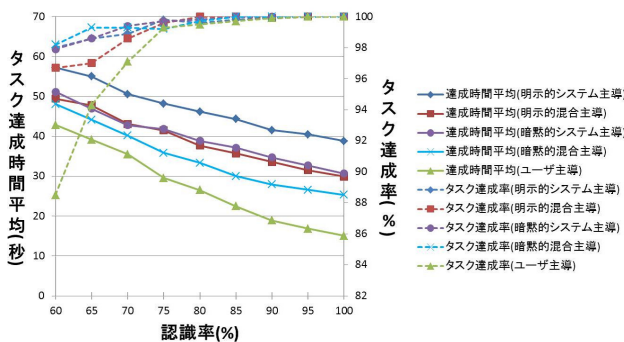


図 3 達成時間と達成率 (余裕時発話条件)

Fig. 3 Completion time and completion rate of condition systems speak users only when users afford speak.

の認識率を上げることはもちろんであるが、ユーザの認識率を 100%にする必要があるということがわかる。

また、余裕時発話条件 (ユーザ認識率 100%) のシステム認識率 80%のときと、ユーザ誤認識条件 (ユーザ認識率 80%) のシステム認識率 100%のときの、ターン数平均を比較すると、システム主導は同程度となる。一方、ユーザ主導と混合主導はユーザ誤認識条件の方が、ターン数平均は少ない。従って、ユーザの認識率が下がることよりもシステムの認識率が下がることの方が問題である。

一般に、車内では雑音が多く、システムの認識率を 100%にすることは容易ではない。それに比べて、ユーザの認識率を 100%近くにするには方法次第では可能である。例えば、余裕時発話条件のように、余裕がある場合のみシステムが発話することでユーザが対話に集中できるので、ユーザの認識率が上がり、ユーザ誤認識条件のようなタスク達成率の低下は起こりにくいと考えられる。これらのことから、余裕時発話条件はユーザの認識率を上げることができると考えられる。

## 4. 主観評価実験

### 4.1 実験条件

自動車を運転する際、運転者には運転のための認知負荷

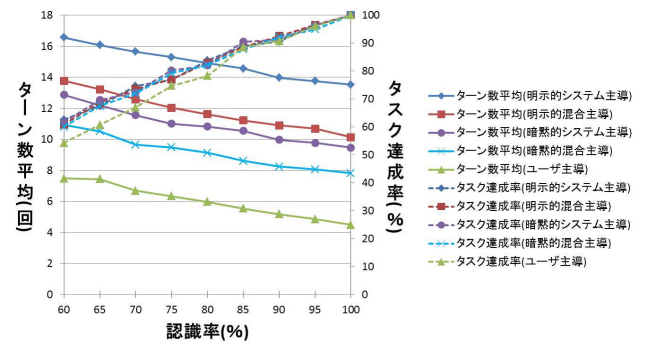


図 4 ターン数と達成率 (ユーザ誤認識条件)

Fig. 4 Turn and completion rate of users misrecognition condition.

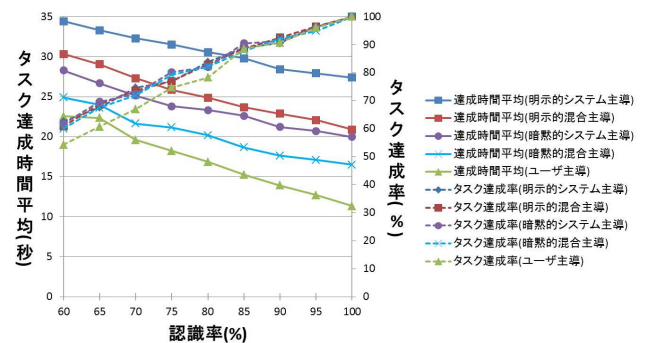


図 5 達成時間と達成率 (ユーザ誤認識条件)

Fig. 5 Completion time and completion rate of users misrecognition condition.

がかかっており、運転中に音声対話を行うときは負荷がかかった状態で使用されると考えられる。そのため本実験では、音声対話を行う第 1 課題を、ボール型ラジコンを操作する第 2 課題と同時に行う、2 重課題法 [4],[5] で実験を行った。実験場所は岡山大学工学部 4 号館 601 室で、被験者は 4 名である。本実験で使用したボール型ラジコンは Orbotix 社製の Sphero 2.0 である (図 6)。Sphero 2.0 は iOS や Android 端末で Bluetooth 接続をして操作が可能である。Sphero 2.0 の仕様を表 1 に示す。今回は最高速度が 0.5 m/s になるように設定した。走行コースは広い床に、底辺の一辺約 6 cm の直方体の障害物を一直線上に 50 cm 間隔で 5 本固定し、スタート地点とゴール地点はそれぞれ障害物の端から 50 cm 離れた場所とする。スタート地点からゴール地点まで、行きは障害物の間をジグザグ走行、帰りは並べられたペットボトルの横を直線で走行する。このコースを 1 回タスクが達成されるまで何周も行う。実験風景と見取り図を図 7 に示す。本実験の流れは以下のようになる。

- (1) 10 分間練習
- (2) 対話戦略の説明 (どうい対話を行うか)
- (3) 想定認識率 70%で音声対話+ラジコン操作 (3 回)
- (4) 評価
- (5) 想定認識率 90%で音声対話+ラジコン操作 (3 回)
- (6) 評価



図 6 ボール型ラジコン (Sphero 2.0) の外観

Fig. 6 Exterior of a radio-controlled ball(Sphero 2.0).

表 1 ボール型ラジコン (Sphero 2.0) の仕様

Table 1 Specification of a radio-controlled ball(Sphero 2.0).

仕様	
直径	74 mm
重さ	168 g
最高速度	秒速 2m
最大通信距離	30 m

(7) (2) から (6) を全ての対話戦略で行う

(3) と (5) は音声対話が始まってからスタートするが、1 回タスクが達成されるたびにコースアウトしてスタート地点まで戻ってもらった。

合成音の作成には NTT 研究所で開発された音声デザインツール Sesign[6] を用いた。音声認識は機械のふりをした人間が行い、想定単語認識率は 70% と 90% の 2 種類で実施した。これは前節の結果より、ユーザ主導では認識率 75% 以下でターン数やタスク達成時間が大きく増加したため、その両側で差のある 70% と 90% を採用した。今回の実験で使用対話戦略を以下に示す。

(1) システム主導

(a) 明示的確認発話

(b) 暗黙的確認発話

(2) ユーザ主導

(3) 混合主導

(a) 明示的確認発話

(4) 余裕があるときのみ発話

(a) 明示的確認発話

余裕があるときのみ発話する場合は、システム主導の明示的確認発話を使用し、帰りの直線のときのみシステムから応答が返ってくるようにした。また、混合主導は他の対話戦略との違いを明確にするため、1 回目は目標物から入力、2 回目は市名から入力、3 回目は県名と市名から入力するようにユーザに指示を行った。ただし、評価の際には混合主導では本来は何から入力してもよいということを考慮するように指示を行った。

#### 4.2 評価尺度

本実験では以下の 3 つの評価尺度に基づき 5 段階で主観評価を行ってもらった。

(1) 効率性

3 回タスクを達成したあとに、行った対話の効率性を

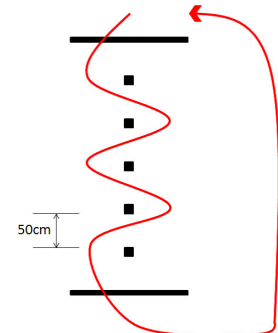


図 7 実験風景-見取り図

Fig. 7 Experiment image-sketch map.

5 段階 (とても効率が良い : 5, 効率が良い : 4, どちらとも言えない : 3, 効率が悪い : 2, とても効率が悪い : 1) で評価。

(2) 非負担度

行った対話が Sphero 2.0 の運転に与えた負担を 5 段階 (とても負担を感じた : 5, 負担を感じた : 4, どちらとも言えない : 3, ほとんど負担を感じなかった : 2, 全く負担を感じなかった : 1) で評価してもらい、その値を 6 から減じた評価値を非負担度とする。

(3) 満足度

行った対話戦略を実際に使用したいかどうかを 5 段階 (とても使いたい : 5, 使いたい : 4, どちらとも言えない : 3, あまり使いたくない : 2, 絶対に使いたくない : 1) で評価。

#### 4.3 実験結果

想定単語認識率ごとの評価値を図 8 と図 9 に示す。3 つの評価値全てについて、認識率が 70% のときの評価値よりも 90% のときの評価値の方が高い。ユーザ主導は認識率 70% では全体的に評価が低く、特に非負担度が低い値となっているが、認識率が 90% になると評価値は高くなる。また、混合主導は認識率によらず、全ての評価値が高くなっている。これは、一発話目が何から入力してもよいという自由度の高さと、認識率の低下がターン数やタスク達成時間に影響を与えないという混合主導の特徴に起因していると考えられる。余裕があるときのみ発話する場合は、非負担度は他の戦略と比べて最も高くなっており、音声対話が運転中にほとんど妨げにならなかったと言える。しかし、効率性や満足度が低いことから、システムの応答がすぐに返ってこないために、効率が悪いとユーザに感じさせてしまったと考えられる。システム主導において、認識率が 70% では、明示的確認発話よりも暗黙的確認発話の方が効率性と非負担度は高く、満足度は低い。これは、暗黙的確認発話の方が明示的確認発話よりもターン数は少ないが、聞き取るべき情報量が多く負担を感じたためだと考えられる。一方、認識率が 90% では、効率性は暗黙的確認発話の方が高いが、非負担度と満足度はほぼ同じ評価値となっている。これは、認識率が 90% だとシステムの誤認識



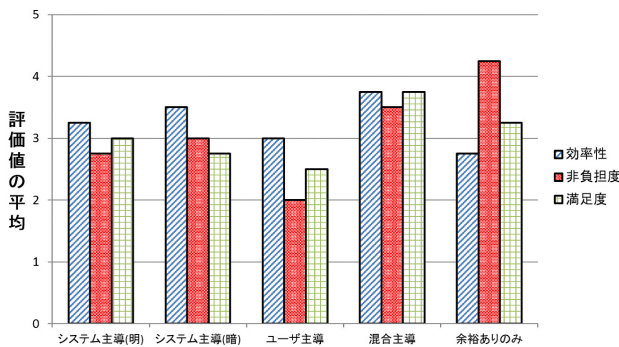


図 8 想定単語認識率 70%

Fig. 8 Assumption word recognition rate 70%.

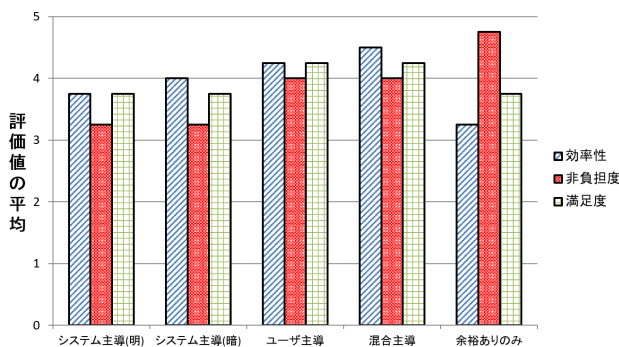


図 9 想定単語認識率 90%

Fig. 9 Assumption word recognition rate 90%.

が少なく、ユーザが明示的確認発話と暗黙的確認発話の違いをあまり感じなかったためだと考えられる。

また、対話戦略によってペットボトルの衝突回数に差が出るのが考えられたが、今回の実験では差は見られなかった。この原因として、本実験で行った対話の負荷では、静止している物体を避ける運転は容易だったと推測できる。

## 5. おわりに

本報告では、安全運転に配慮した音声対話戦略を検討するために、異なる対話戦略に対する客観的評価と主観的評価を行った。まず、客観評価実験として計算機シミュレーションを行い、ターン数平均、タスク達成時間平均とタスク達成率の評価を行った。客観評価の結果では、ユーザ主導の対話戦略は認識率が高いときはターン数やタスク達成時間は少ないが、認識率が低くなるとターン数やタスク達成時間が増えてしまうということ、システム主導や混合主導はターン数やタスク達成時間が増えてしまうが、認識率が低くなってもターン数やタスク達成時間の大幅な増加はないということが示された。よって、システムの認識率が高いと仮定できる場合はユーザ主導が優れているが、システムの認識率が低いと仮定される場合はシステム主導や混合主導の対話戦略が優れている。確認発話の種類では、ターン数やタスク達成時間が少なくなることから、明示的確認発話よりも暗黙的確認発話の方が、有効な手法である。また、運転に集中しているために、ユーザの認識率が 80%に

下がってしまうと仮定してシミュレーションを行った結果、タスク達成率は大幅に低下した。ユーザ認識率を 100%に近づけられるように、余裕があるときのみ発話するという戦略は運転中の対話戦略として適していると言える。

続いて、主観評価実験としてラジコンを操作しながら音声対話を行う 2 重課題法を行い、効率性、非負担度と満足度の評価を行った。主観評価の結果、ユーザ主導は認識率 70%では評価は低いですが、認識率が 90%になれば高い評価が得られた。この結果は客観評価の結果と一致しており、ユーザ主導は認識率が高いとターン数やタスク達成時間が少ないため好まれるが、認識率が低いとターン数やタスク達成時間が増えて負担が大きくなるため、ユーザは煩わしく感じてしまう。また、混合主導では認識率によらず評価が高かった。これは他の手法よりも自由度が高く、認識率が低くてもターン数やタスク達成時間の増加量が少ないためだと考えられる。余裕があるときのみ発話する場合は、効率性と満足度は低いですが、非負担度は他の戦略と比べて最も評価が高い。これは、タスク達成までに時間がかかるため効率が悪いが、ユーザが感じる負担はほとんどなかったためだと考えられる。

今後の課題として、対話戦略間でペットボトルとの衝突回数に差が見られなかったことの原因調査が挙げられる。この原因は、本実験で利用した音声対話タスクの負荷があったとしても、静止している物体を避けることは簡単だったためだと考えられる。実際の道路で運転する場合は動いている物体を瞬時に認識しなければならない。従って、コース上に動く障害物を取り入れることで、対話戦略ごとに衝突回数の差が現れると考えられる。

謝辞 本研究は公益財団法人岡山工学振興会の助成を受けたものです。

## 参考文献

- [1] 千葉成美: “第 13 編・第 19 部門・第 3 章 音声認識合成応用システム,” 電子情報通信ハンドブック (第 1 分冊), 電子情報通信学会 (編), pp. 1204–1205, オーム社, 1988.
- [2] 西本卓也: “音声インタフェースは本当に人に優しいか?,” 人工知能学会研究会資料, SIG-SLUD-A202-05, pp. 27–32, Nov. 2002.
- [3] 石川泰: “音声対話システムの評価法,” 日本音響学会誌, vol. 54, no. 11, pp. 807–811, Nov. 1998.
- [4] 脇田敏裕, 寺島立太, 小島真一, 清水司, 本郷武朗: “運転中情報機器操作性の評価法,” 情報処理学会論文誌, vol. 42, no. 7, pp. 1762–1769, 2001.
- [5] 西本卓也, 高山元希, 櫻井晴章, 荒木雅弘: “音声インタフェースのための対話負荷測定法,” 電子情報通信学会論文誌, vol. J87-D-II, no. 2, pp. 513–520, 2004.
- [6] 阿部匡伸, 水野秀之, 水野理, 野田喜昭, 高橋敏, 中嶋信弥: “音声デザインツール Sesign,” 電子情報通信学会論文誌, vol. J84-D-II, no. 6, pp. 927–935, 2001.