

# 車載情報サービス用音声対話コーパス

## A SPOKEN DIALOG CORPUS FOR CAR TELEMATICS SERVICES

立石 雅彦\*1 赤堀 一郎\*1

Masahiko Tateishi Ichiro Akahori

スコット ジュディ\*2 大淵 康成\*3 三田村 照子\*2 エリック ナイバーグ\*2 畑岡 信夫\*4

Scott Judy Yasunari Obuchi Teruko Mitamura Eric Nyberg Nobuo Hataoka

### 1. まえがき

近年、ドライバーが車の中に居ながらにしてインターネット等のネットワークにアクセスし、交通情報やニュース等のオンライン情報を入手する車載情報サービスが立ち上がりつつある。

車載情報サービスが提供する豊富かつ多様な情報を自由に検索する手段として音声対話が注目されている<sup>8),9)</sup>。しかし車載端末は計算資源に限られており、情報センター間との通信回線も決して十分な帯域を持っていないことから、車載端末と情報センターが協調して音声対話を効率的に進める仕組みが必要となる。

本稿は我々が提案するクライアント - サーバー型音声対話システム Conversational Agent for Multimedia Mobile Information Access について述べる。CAMMIA は VoiceXML<sup>7)</sup> を拡張した新規対話シナリオ記述言語を使用し、柔軟な話題遷移を可能とする。この特長により、ドライバーは目的地設定対話の途中で目的地の天気を問い合わせ、また目的地設定対話に戻るといった対話を行うことができる。さらに音声対話の大部分を車載端末で処理することで通信回数を低減できる。以下、CAMMIA の構成と動作、音声認識用文法の設計、CAMMIA 設計 / 評価用に収集した音声対話コーパスについて述べる。

### 2. CAMMIA のシステム構成

CAMMIA のシステム構成を図 1 に示す。車載端末は VoiceXML インタプリタ、音声認識システム、音声合成システムを備える。情報センターは対話管理部、情報データベース、対話シナリオを備える。

情報センターの対話管理部は車載端末から要求を受けると、経路案内、天気案内等の話題単位の対話シナリオを動的に生成する。そして該対話シナリオとその対話に必要な音声認識語彙・文法をまとめて車載端末に配信する。車載端末は該対話シナリオを VoiceXML インタプリタにより解釈実行し、ドライバーとの音声対話を行う。この間、情報センターとの通信は発生しない。また、該対話シナリオには話題遷移を検出する機能が組み込まれている。車載端末は話題遷移を検出すると新しい話題の対話シナリオを情報センターに要求 / 受信し、該対話シナリオを用いて音声対話を継続する。

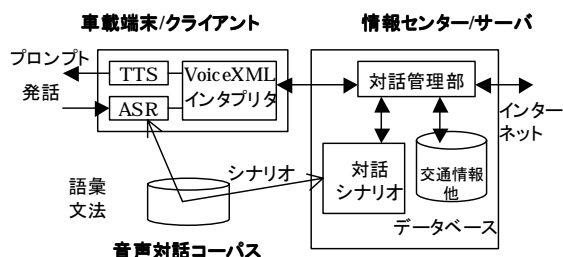


図 1: CAMMIA のシステム構成

音声対話コーパスは音声認識用文法、および対話シナリオの設計 / 評価に使用する。以下、音声対話コーパスの利用例として音声認識用文法の設計を簡単に説明した後、音声対話コーパスについて述べる

### 3. 音声認識用文法の設計

車載端末の限られた計算資源で音声対話を実現するため、CAMMIA では場面に応じた複数の小規模文法を設け、必要に応じ情報センターから車載端末に配信する。文法設計を効率化するため、我々は音声対話コーパスを用いて文法を自動生成する方法を取った。最初に設計者が単一文法により文法を記述する。これは正規文法に比べ様々な制約を効率的に記述できる。次にこれら制約を全て品詞と見なすことで文脈自由文法に変換する。この文法を再帰回数に制約を設けて正規文法に変換する。一般にこうして得た文法は冗長である。そこで、対象となる場面に関連する文を集めたコーパスを用意し、全文を上記文法で解析する。その後、一度も使われなかった状態とアークを除去し、文法を簡略化する (図 2)。

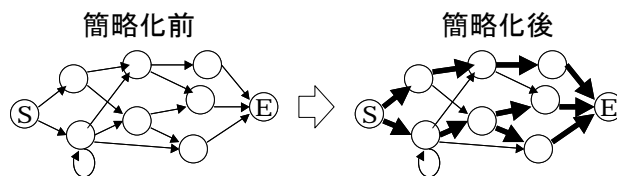


図 2: 音声対話コーパスによる正規文法の簡略化

\*1 デンソー基礎研究所

\*2 カーネギーメロン大学 Language Technologies Institute

\*3 日立製作所基礎研究所

\*4 日立製作所中央研究所

## 4. 音声対話コーパスの作成

本節では、音声対話コーパスの話者の選定と収集期間、収集方法について説明する。

### 4.1 話者の選定と収集期間

話者の発話に出現する語彙や文法、および音声対話に出現する話題は話者の年齢や性別に大きく左右されると考えられる。したがって幅広い年代の男女コーザに対応できる音声対話システムを構築するには、収集する音声対話コーパスの話者の性別と年代を均等化することが望ましい。そこで我々は、男性話者 137 名、女性話者 113 名の計 250 名を話者とした。話者の年代も 20 代、30 代、40 代、50 代、60 代でほぼ均等とした。話者は首都圏在住者で、うち 235 名は運転免許所有者であった。また、50 名はナビゲーションの使用経験者であった。

次に我々は、上記 250 名の話者を 50 名ずつの 5 つのグループに分けた。そして各グループの音声対話を 1 ヶ月から 1 週間程度の時間間隔を空けて順に収集した。1 グループの収集が完了した時点で、音声対話収集方法を改善し、次の収集に反映した。全グループの音声対話収集には約 4 ヶ月を要した。

### 4.2 音声対話収集方法

本節では音声対話収集に用いた音声対話収録環境、及びタスクと教示について述べる。前節で述べたように、我々は 50 名の音声対話を収集する度に音声対話収集方法を改善した。我々は最初の 50 名の音声対話収集後に、最も大きな改善を行った。以下、その改善前、改善後の音声対話収集方法を説明する。

#### 4.2.1 音声対話収集方法（改善前）

##### ・音声対話収録環境

我々は本音声対話を東京都内のスタジオで収録した。該スタジオは 2 部屋で構成され、片方の部屋にドライブ情報を問い合わせる話者が、他方の部屋にドライブ情報を提供するオペレータが入った。話者の部屋は車載端末を備えた車に相当し、オペレータの部屋は情報センターに相当する。また、話者はドライバーに相当し、オペレータは音声対話システムに相当する。話者とオペレータはヘッドセットとマイクを使用して音声対話を行う。オペレータは以下の情報をもつサービス要員として、話者の質問に回答した。

- オンライン交通情報
- 飲食店情報
- 観光情報
- ホテル、旅館情報

オペレータには、正確な日本語を話すだけでなく、旅行案内という専門性の高いタスクを巧みにこなす能力が求められる。そこで我々は旅行代理店および民放ラジオ局の勤務経験があり、かつ旅行主任技術者の資格を有する熟練アナウンサーを起用した。

##### ・タスクと教示

我々は話者に対し、車載端末を搭載した車に乗車し、東京から伊豆または箱根に旅行するという想定で、オペレータと対話し、ドライブ情報を得よう指示した。

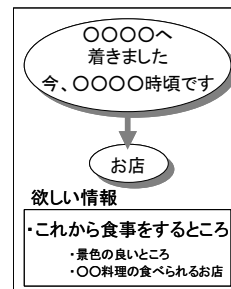


図 3：タスク教示書の例

本音声対話コーパス収集に先立ち、302 名のドライバーに対して行った事前調査では、運転中に欲しいと思った案内の上位 3 つは交通案内、飲食店案内、観光案内であるという結果が得られていた。したがって、我々は最初に以下の 3 つのタスクを話者に与えて音声対話コーパスを収集した。

- タスク 1. 目的地までの交通情報を得る
- タスク 2. 目的地付近で昼食を食べる飲食店を探す
- タスク 3. 昼食後に観光する場所を探す

タスク教示は書面を用いて行い、状況と指示をそれぞれ文章で記述した。図 3 にタスク 2 の教示書の例を示す。

#### 4.2.2 音声対話収集方法（改善後）

音声対話収集方法の改善内容を説明する。はじめにタスクと教示の改善について説明し、その後、音声対話収録環境の改善について述べる。

##### ・タスクと教示

前節のようにあらかじめ決められたタスクを被験者に単に与える方法は幾つか問題点のあることが判明した。最も大きな問題は、話者の動機付けが十分にできないことである。被験者は旅行中であると十分思い込めず、与えられたタスクに関する質問ができない状況が発生した。その場合、オペレータは対話を中断し、どのような質問が可能なのか話者に教示し、場合によっては最初から対話をやり直す必要が生じた。また、教示を図 3 のように文章で記述すると、多くの話者は対話開始時に教示書を読んでしまい、自発発話の収集に悪影響を与えることが判明した。以下に教示書の影響（太字）を受けた対話例を示す。

話者： **山中湖へ着きました**  
オペレータ： はい  
話者： **えー今 12時半ごろです**  
オペレータ： はい

そこで我々は残りの話者 200 名に対し、以下のように教示方法を変更した。200 名を 5 名のメンバーからなる 40 のグループに分けた。そして各グループに伊豆、箱根いずれかを目的地に選ばせ、その目的地に一泊二日のドライブ旅行をするという前提で議論をさせ、自由に計画を立てさせた。議論の後、各話者に各自の興味や関心事を反映した 2 つのタスクセット A、B を作成させた。タスクセット A は旅行 1 日目に自宅から旅行に出発する際に知りたい情報を列挙したりリスト、タスクセット B は旅行 2 日目にホテル

や旅館をチェックアウトする時点に知りたい情報を列挙したリストである。さらに、各話者の音声対話収録を2回に分け、最初の収録でタスクセットAに関する収録を行い、1時間ほど時間を置いた後にタスクセットBに関する収録を行った。

#### ・ 音声対話収録環境

音声対話収録時には、旅行の日にちを事前に条件設定しておかなければならないことが分かった。日にち設定は交通案内、ガソリンスタンド案内等に影響する。たとえば、交通状況は平日か週末かで混雑する道路や混雑の程度が大きく変化するので、オペレータの案内もそれに依りて変えなければならない。

しかし日にちを条件設定していないと、音声対話収録中に「今日は何曜日でしたっけ・・・」といった条件を確認する対話が入ってしまう。そこで、各被験者に対し、音声対話収録前に、旅行に出発する日を平日にするか週末にするかを選択させ、それに従い日にちを決定した。

これら日にち等の条件設定についても、書面に条件を文章で記述すると話者とその文章を読み上げてしまう可能性がある。そこで我々は、話者の机上に卓上カレンダー、置時計を置いて日にちや時刻を示す等、条件設定を物で示す工夫をした。

また、次に収録を行う話者（以下、次話者）の緊張をほぐし、収録環境やタスクに慣れさせるため、ドライバー役の話者の隣に同乗者役として座らせた。次話者はヘッドセットを着用し、オペレータと話者の音声対話を聴くことができるが、音声対話に加わることはできない。

次話者を同乗者役として収録に立ち合わせたところ、次話者をタスクに慣れさせる効果だけでなく、以下のように話者の質問を引き出す効果もあることが分かった。

- (i) タスクセット中の項目で、話者が質問を忘れていた項目があると、次話者がその項目を示して質問を促す。
- (ii) オペレータが「（お問い合わせの情報を）お調べいたしますのでしばらくお待ちください」と対話を中断したとき、話者と次話者が会話して質問項目の確認などをする。

例) 話者と次話者の会話（太字の箇所）

オペレータ： 箱根湯本からえー芦の湖畔までの時間と  
いうことでお調べいたしますので

話者： はい

- ・ **お店の名前まで聞いていいのかな**
- ・ **うんいんじゃない**
- ・ **実際にあるお店でも**
- ・ **いんじゃないいんじゃない**

オペレータ： はい大変お待たせいたしましたえーまず  
中目黒から

以上の改善の効果についてオペレータに尋ねたところ、改善後の話者は積極的に質問をするようになったと意見を述べた。また、オペレータが対話を中断し話者に教示を与えた回数は改善前では50名の話者に対して68回だったのに対し、改善後の200名の話者に対してはわずか1回と大幅に削減できた。

## 5. 音声対話コーパスの分析

### 5.1 話者発話中の語彙の分析

本音声対話コーパスは450の対話からなり、オペレータ、話者それぞれの発話数は34,051発話、33,072発話であった。

話者の語彙数は4,603個であり、内999個が固有名詞、残りの3,604個がそれ以外の単語であった。これより本音声対話コーパス中の話者の発話認識に必要な音声認識システムの語彙数は5,000個程度であることが分かった。

### 5.2 音声対話コーパス中の話題の分析

本音声対話コーパス中に出現した話題数は話者250名で1253個、話者1名あたり平均5.0個となった。出現話題は上位から順に交通案内(28%)、飲食店案内(21%)、観光施設案内(15%)、景勝地案内(15%)となった。音声対話コーパス中の話題の例えば90%以上をカバーするには上位7種類の話題に対応しなければならないことが分かった。

## 6. まとめと今後の課題

車載情報サービス用の音声対話コーパスを250名の話者から収集した。話者の発話の語彙数は4,603個であった。また、音声対話コーパス中の話題の例えば90%以上をカバーするには上位7種類の話題に対応する必要があることが分かった。現在、これらの結果を元に音声認識システムの語彙・文法、および対話シナリオの設計を行っている。

### 謝辞

本音声対話コーパス開発にあたり貴重なご意見、ご指導をいただいた京都工芸繊維大学、荒木雅弘助教授に感謝します。また、本音声対話コーパスの分析に関し、貴重な助言をいただいた日立中央研究所、渡邊純一郎氏に感謝します。

### 参考文献

- 1) E.Nyberg et al., DialogXML: Extending VoiceHML for Dynamic Dialog Management, Proc. of HLT-2002(2002)
- 2) 畑岡他, VoiceXML をベースにした音声対話方式の開発, 日本音響学会 2002 年秋季研究発表会 (2002)
- 3) Y.Obuchi et al., Robust Dialog Management Architecture Using VoiceXML for Car Telematics Systems, Proc. of Workshop on DSP in Vehicular and Mobile Systems(2003)
- 4) M.Tateishi et al., A Spoken Dialog Corpus for Car Telematics Services, Proc. of Workshop on DSP in Vehicular and Mobile Systems(2003)
- 5) 立石他, 車載情報サービス用音声対話コーパス, 情報処理学会研究報告, 2003-SLP-46, pp.43-48(2003)
- 6) 大淵他, VoiceXML をベースにした頑強な音声対話管理アーキテクチャ, 情報処理学会研究報告, 2003-SLP-46, pp.49-54(2003)
- 7) VoiceXML forum, <http://www.voicexml.org/>
- 8) 伊藤他, ドライブプランニングシステムの自然言語インタフェース, 人工知能学会誌, Vol.17, No.3, pp.285-290(2002)
- 9) 河口他, 実走行車内における音声データベースの構築, 情報処理学会研究会, 音声言語情報処理, 99-SLP-30-12(2000)