

河口 信夫^{†1†2}松原 茂樹^{†1†3}武田 一哉^{†1†2}外山 勝彦^{†1†2}稲垣 康善^{†2}

†1 名古屋大学統合音響情報研究拠点 (CIAIR)

†2 名古屋大学大学院工学研究科 †3 名古屋大学言語文化学部

E-mail: kawaguti@nuie.nagoya-u.ac.jp

1 はじめに

近年、高度道路交通システム (ITS: Intelligent Transport Systems) の実現のために、様々な研究開発が進められている。その中でも、車室内における高度情報システムは、ドライバーとのフロントエンドとして高い期待が寄せられている [1, 2]。実走行車内において操作可能な情報システムを実現するためには、高騒音下でのロボスト音声認識技術と、自然発話を受理可能な音声対話技術が必要である。名古屋大学統合音響情報研究拠点 (CIAIR) では、ロボストな車内音声対話システムの実現を一つの目標と定め、その要素技術の研究開発を進めている。

音声言語処理の発展のために、音声コーパスは重要な役割を果たしてきた [3]。しかし、従来の音声コーパスは防音室や静かな室内で収録された音声を中心であり、実環境における大量データの収録は行われていない。また、車両運転中にドライバーがシステムとの対話を行う場合、運転中に特有な言語現象が現れたり、運転操作と関連した発話が行われたりする可能性が高く、模擬環境を用いて、このような対話を収録することは困難である。

CIAIR では、このような実走行車内環境での音声処理技術の研究開発を促進するために、大規模な車内音声データベース、および車内音声対話コーパスの構築を行うこととした [4]。本稿では、CIAIR で構築するデータベースの概要について述べ、特に音声対話コーパス構築の目的、収集方法、収集状況について述べる。

2 車内音声データベースの概要

音声データベースとしては、ドライバーを被験者として、実走行中の車内において、音素バランス文、連続数字、および離散単語の読み上げを行う。また、自然発話を受理可能な音声対話技術の獲得のために、ナビゲーションやレストラン等の情報検索が可能な車内情報システムとドライバーとの間の対話を収録し、書き起しによって対話コーパスを構築する。

今回の収録では、車内情報システムを利用している環境を模擬的に実現するために、制限された発話を行うように訓練されたナビゲータが同乗し、情報システムの役割を果

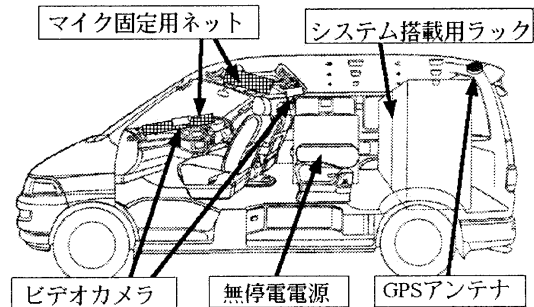


図 1: 収集用車両の概要

たしている。将来的には、ナビゲータを Wizard of OZ のシステムや実際の対話システムと置き換えて収録していくことを計画している。

また、収録するデータは、音声のみではなく、画像やハンドル操舵角力、アクセル・ブレーキの踏力、GPS による車両位置が同期されたマルチモーダルデータであり、今後の様々な研究に活用できることを配慮している。本データベースの収録のために、複数のモダリティの同期入力を可能とする専用のデータ収録車 (図 1) を作成した。

3 車内音声対話コーパス構築の目的

走行車内における音声対話コーパスの構築は、車内対話システムの基本資料収集を目的としている。本コーパスにより、走行車内特有な言語現象、言い回し、語彙、発話の重なり具合の調査・分析が可能となり、車内対話をモデル化することができる。また、ドライバーやシステムからの発話のタイミングを調査することにより、漸進的な音声言語処理の基礎データが得られる。走行車内のヒューマンインタフェースの研究データとしては、運転歴、運転状況の違いによる理解能力、発話内容、発話単位、発話速度の変化等を得られる。

単なる対話のみではなく、画像や車両操作による運転状況のマルチモーダル情報が統合的に記録されたデータが得られるため、より高度なインタフェース構築の基礎データとして活用できる。また、車内対話タスクは、ナビゲーション、情報検索、車載機器の操作等の複数のドメインを同時に扱うため、マルチドメインシステム [5] 構築の基礎データとなっている。

今回の収録では、ナビゲータが機械の代役を担っており、制限されたナビゲータによる発話は、近未来の車内音声対話システムの方向性を示唆するものとして利用できる。

Construction of Spoken Dialogue Corpus in Moving Car

Nobuo Kawaguchi^{†1†2}, Shigeki Matsubara^{†1†3}, Kazuya Takeda^{†1†2}, Katsuhiko Toyama^{†1†2} and Yasuyoshi Inagaki^{†2}

†1Center for Integrated Acoustic Information Research (CIAIR), Nagoya University. †2Graduate School of Engineering, Nagoya University. †3Faculty of Language and Culture, Nagoya University.

すこし、喉が乾きました。運転にちょっと
疲れたので、一休みしたいですね。

今日は、彼（彼女）の誕生日！
これから誕生日祝いをしたいと思います。

図 2: 状況プレートの例

4 車内音声対話コーパスの収録手法

実際の走行車内で車内情報システムを利用する環境を想定し、1時間強の走行コースを定め、仮想目的地を目指すナビゲーションと、レストランや駐車場などの情報検索や予約を行う情報検索タスクを行い、音声対話の収録を行っている。今回の収録では、車内情報システムを模擬的に実現するために、発話を制限するように訓練されたナビゲータが同乗して情報システムの役割を果たした。ナビゲータは事前に走行コースを4~8回程度試走し、ナビゲーションに対する習熟を行っている。ナビゲーションは基本的に、車両位置に対するイベントであり、車両が特定の場所に来た時に、特定の指示を行う。また、ドライバからの問いかけに対しては、随時応答を行う。また、ナビゲータ用に想定対話などを含むマニュアルを作成した。この中では、ナビゲータには視覚情報やジェスチャによる応答を行わないように指示してある。

4.1 指示・状況プレート

情報検索タスクに対しては、事前に指示を与えても、運転中にドライバがそれらを思い出して実行することは困難である。そこで、ドライバに対する検索タスクの指示のために「指示プレート」を用いる。指示プレートには「和食」「中華」「洋食」「喫茶店」「コンビニ」「銀行」「駐車場」といった短い単語が書かれている。ドライバは、ナビゲータからプレートを見せられた時、その単語に関する要求発話を行うように指示されている。

また、自然発話を促し、対話に広がりを持たせるために、状況を表す文「状況プレート」を作成し、文章をそのまま読み上げても、情報システムに対する指示にならないような工夫を行なった。状況プレートの例を図2に示す。短い文であるが、走行中に読むことは危険であるので、信号待ち等の余裕がある時にドライバに見せることとした。また、そのまま読み上げることを防ぐために、プレートを見せた直後には発話せず、プレートを隠した後に、発話することとした。

5 収集データの状況

図3に、収録した対話の一部を書き起こしたものを示す。この対話では、図2の最初の状況プレートの読み上げに近い発話も生じている。ただし、このような発話はタスクの最初の発話に特徴的であり、以後の対話は自然発話が中心になっている。また、04 N, 05 Nに見られるように、ナビゲータはドライバとの対話を行いながら、同時にナビゲーションを行っており、喫茶店を探すタスクの途中にナビゲーションタスクが割り込んでいる状況が発生している。

6 まとめ

CIAIRでは、ロバストな音声対話認識技術確立のために、高騒音下である実走行車室内の音声データベースおよび、音声対話コーパスの構築をはじめている。タグづけに関しては、音声対話コーパスの共有化[6]のために、共有可能な

- 01 D: ちょっとのどが渴いてきたぞ お茶でも飲んでいき
いんだけど
02 N: はい 喫茶店でいいですか
03 D: はい
04 N: はい では喫茶店を
05 N: 次の信号を左折してください
06 D: はい
07 N: この近くに喫茶店は4軒あります
08 E: <ウインカー>
09 D: アイスコーヒーの美味しいところがいいですね
10 N: アイスコーヒーの美味しいのはアトリエシュル
プランというお店がお勧めです
11 D: はい コーヒーの専門店?
12 N: はい コーヒーの専門店です
13 D: はい
14 N: そちらでいいですか
15 D: いいです この時間空いていますか
16 N: えと この時間は空いています
17 D: はい
18 N: はい
19 D: ではそこに行きましょう
20 D: その喫茶店は駐車場ありますか
21 N: はい 駐車場は無料駐車場が2台あります
22 D: はい
23 N: では駐車場をご案内します
24 D: はい
25 N: 次の信号を左折してください
26 D: はい

図 3: 車内対話の例 (D: ドライバ, N: ナビゲータ)

タグ構造を採用することを予定している。また、Wizard of OZ 法や音声認識を用いた対話システムを用いた音声対話の収録を進めていく計画である。収録した音声や対話データは、整理を行って随時公開していく予定である。本車内音声対話コーパスは構築を始めたばかりである。意見、要望等があれば、忌憚なくお知らせ願いたい。

謝辞

本研究は文部省科学研究費補助金 COE 形成基礎研究費 (課題番号 11CE2005) の補助を受けて行われた。データベース収集の目的や方法について熱心に検討を頂いた CIAIR コーパス WG の諸氏に感謝致します。

参考文献

- [1] 松原茂樹, 河口信夫, 外山勝彦, 稲垣康善: 音声対話インタフェースを備えた車内秘書システムの提案, 人工知能学会全国大会 (第 13 回) 論文集, 16-02, pp.231-234(1999).
- [2] 松原茂樹, 河口信夫, 外山勝彦, 稲垣康善: 発話の同時理解・同時生成に基づく車内音声対話秘書システムの提案, 人工知能学会研究会, 話し言葉と理解, SIG-SLUD-9902, pp.1-6(1999).
- [3] 板橋秀一: 音声コーパス, 情報処理, Vol.38, No. 11, pp.1012-1018(1997).
- [4] 河口信夫, 松原茂樹, 岩博之, 梶田将司, 武田一哉, 板倉文忠: 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).
- [5] Nobuo Kawaguchi, Shigeki Matsubara, Katsuhiko Toyama, Yasuyoshi Inagaki: An Architecture for Multi-Domain Spoken Dialog Systems, Proceedings of the 5th Natural Language Processing Pacific Rim Symposium(NLPRS'99), pp.463-466(1999).
- [6] 土屋 俊, 堀内靖雄, 石崎雅人, 前川喜久雄: 音声対話コーパスの共有化へ向けて, 人工知能学会誌, Vol.14, No.2, pp.231-242(1999).