

ユーザの発話傾向分析に基づく 車載機器操作のための音声入力手法の検討

倉田 岳人^{†1} 市川 治^{†1} 西村 雅史^{†1}

ドライバーへの負荷を軽減する、車載機器操作手段の一つとして音声認識が取り入れられてきた。しかし、あらかじめ定められた定型句からなるコマンドしか受理できない従来の音声コマンド方式では、ユーザが前もってコマンドを正確に記憶していることを前提としており、音声入力に不慣れたユーザが直観的に音声だけで車載機器を操作することは困難である。本報告では、車載機器操作のための発話を多数のユーザから収集し、それらに分析を加え、従来方式の限界を定量的・定性的に検証した。さらに、大語彙連続音声認識と自然言語理解の組み合わせにより、ユーザの自由な発話を受理できる拡張音声コマンド方式を開発した。本方式を利用して、実際の車載機器の音声操作機能を実装し、92%を超えるタスク達成率を得ることができた。

Speech Input Method in Automobiles Reflecting Analysis on How Users Speak

GAKUTO KURATA,^{†1} OSAMU ICHIKAWA^{†1}
and MASAFUMI NISHIMURA^{†1}

Automatic Speech Recognition (ASR) has been adapted as a method to control embedded systems in automobiles. The most used approach “Voice Command” is difficult to use for normal users because it accepts only the predefined commands. In this report, we qualitatively and quantitatively show the limitations of “Voice Command” by collecting and analyzing the real utterance examples of users. We also developed “Expanded Voice Command” that can accept users’ unconstrained utterance by combining “Large Vocabulary Continuous Speech Recognition” and “Natural Language Understanding”. We implemented the ASR function of a stock car with “Expanded Voice Command” and confirmed that it achieved more than 92% of task completion ratio.

^{†1} 日本アイ・ピー・エム株式会社 東京基礎研究所
IBM Research - Tokyo

1. はじめに

ドライバーの負荷を軽減し、運転の安全性を向上させるため、車載機器の操作方法の一つとして音声認識が取り入れられて久しい^{1),2)}。一般的に利用されている車載機器操作のための音声認識では、「自宅を目的地に設定する」というような機能に対して、「自宅に帰る」というコマンドが割り当てられている(音声コマンド方式)³⁾。ユーザはこのコマンドを正確に覚えて発話する必要があるが、実際のユーザは事前に定義されるコマンドでは網羅することができないくらい多種多様な発話を行う⁴⁾⁻⁶⁾。この結果として、誤認識、誤動作が生じ、自動車内で音声認識を利用するユーザが増えない原因ともなっている⁷⁾⁻⁹⁾。

誤認識、誤動作を防ぐために、認識対象外の発話を受け付けないようにする方法が検討されてきた^{10),11)}。しかし、これらの手法では、基本的にユーザの多様な発話は認識対象外と判断されてしまう。

ユーザの意図を音声対話により理解する方法も検討されており、実走行中の車内で多数の被験者による音声対話を収録したデータベースも整備されてきた¹²⁾。しかし、安全性の観点から音声対話は多用することができない、との指摘もある^{8),13)}。

一方、我々は、ユーザが車載機器を一発話で音声操作する際の多種多様な表現に対応する方法として、ユーザの発話を書き起こす「大語彙連続音声認識」と、書き起こし結果からユーザが意図した車載機器の機能を判断する「自然言語理解」を組み合わせた音声入力手法を開発した。これを「拡張音声コマンド方式」と呼ぶ。

本報告では、車載機器操作のために行う発話を大量に収集し、それらに対して定量的・定性的な分析を加えた。最初に、従来の音声コマンド方式ではユーザの多様な表現に対応することは難しいことを明らかにする。さらに、音声入力による車載機器操作の性能評価実験を行い、音声コマンド方式に比べ、拡張音声コマンド方式が、幅広い表現を受理でき、高いタスク達成率を得られることを示す。

2. 音声コマンド方式と拡張音声コマンド方式

従来の音声コマンド方式の問題点についてまとめ、拡張音声コマンド方式の仕組みと利点を説明する。

2.1 音声コマンド方式の問題点

従来の音声コマンド方式では、音声操作の対象とする車載機器の各機能に対して、1つ、もしくは複数のコマンドが事前に定義されている。ユーザは、コマンドを発声し、それが正しく認識されると、コマンドに対応する車載機器の機能が起動される。

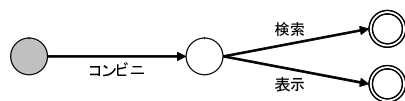


図 1 文法に基づく音声認識
Fig. 1 Speech recognition based on grammar.

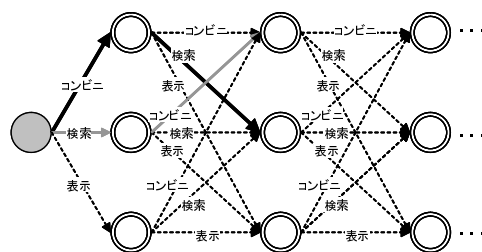


図 2 単語 n -gram に基づく音声認識
Fig. 2 Speech recognition based on word n -gram.

音声コマンド方式では、通常すべてのコマンドを文法で記述し、それに基づき音声認識を行う。図 1 にコマンドを文法で記述した例を示した。文法に基づく音声認識では、左端の灰色の丸を始点とし、右端の二重丸が終点となるような認識結果を、音声と照合しながら探索する。この結果、車載機器の設計段階で各機能に対応付けられ、文法に組み込まれた表現以外は受理されないことになる。つまり、事前にマニュアルを読み、そこに記載されている表現を正確に発話するユーザしか利用することができない。

また、「ユーザが受理されると考えるコマンド」と「音声コマンド方式が受理できるコマンド」が乖離していることも問題となりうる¹⁴⁾。例えば、「コンビニ検索」というコマンドと、図 1 の文法を例に説明する。マニュアルを読んで「コンビニ検索」というコマンドを覚えたユーザであっても、「検索コンビニ」という発話を行うことがある^{*1}。この時、ユーザは正しくコマンドを発話したつもりになっている。しかし、文法に基づく実装を行っている音声認識では通常「コンビニ検索」という表現しか受理できず、ユーザの「検索コンビニ」という発話を正しく認識することはできない。結果として、ユーザは、マニュアルに目を通してコマンドを覚え、正しいコマンドを発話したはずなのに、意図した機能が実行されず、大きなストレスを感じるようになる。このような音声コマンド方式の問題点を克服するためには、正確にコマンドを覚えなくても車載機器を操作できるような音声入力手法が必要である。

2.2 拡張音声コマンド方式の仕組みと利点

拡張音声コマンド方式では、ユーザの自由な発話を大語彙連続音声認識で書き起こし、その書き起こしたテキストに対して自然言語理解を行い、ユーザが意図した機能を判断する。

2.2.1 車載機器操作のための発話を対象とした大語彙連続音声認識

ユーザの自由な発話を書き起こすため、音声コマンド方式で利用するような文法ではなく、単語 n -gram を利用して大語彙連続音声認識を行う^{4),15)}。文法は人手で記述されるた

*1 電話自動応答においても、同様の問題が指摘されている⁶⁾。

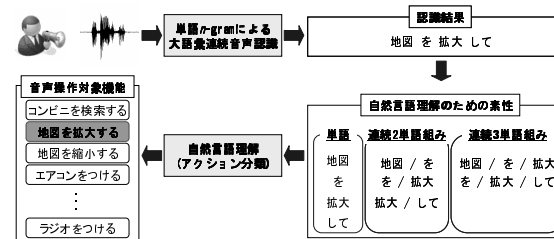


図 3 拡張音声コマンド方式の処理の流れ
Fig. 3 Flow of expanded voice command.

め、受理できる表現に制限があったが、単語 n -gram は大量の学習データから任意の単語間の接続確率が自動的に推定されるため、任意の認識対象単語の接続を受理することが可能となる¹⁶⁾。単純な例として、図 2 に「コンビニ」、「検索」、「表示」の 3 単語のみからなる単語 n -gram を利用した音声認識の概要を示した。単語 n -gram を用いる場合でも、左端の灰色の丸を始点として、音声と照合しながら二重丸で示された終点までの探索を行う。始点以外の任意の状態が二重丸、つまり終点となりうる。すべての矢印、つまり単語接続に対して確率が与えられ、原理的にはあらゆる単語接続を確率的に受理可能となる。例えば、黒太線で示した状態遷移は「コンビニ検索」という発話を受理する。また、文法では受理できなかった「検索コンビニ」という発話も、灰色の状態遷移によって受理される。

2.2.2 認識結果に対する自然言語理解

大語彙連続音声認識による認識結果は自然言語によるテキストで与えられる。これを理解し、ユーザが意図した機能に振り分ける必要がある。自然言語で表現されたユーザの意図の分類は「アクション分類」とも呼ばれ、コールセンタに電話を掛けてきた人を、様々な専門のエージェントに割り振る場面などでも利用されてきた¹⁷⁾。アクション分類の方法としては、1 章で述べたデータベース¹²⁾を利用して事例に基づき分類する方法^{18),19)}や、LSA を利用する方法²⁰⁾などが提案されてきたが、本報告では最大エントロピー法を利用した^{4),21)*2}。

まず、認識結果の単語列を、単語そのもの、連続 2 単語組み、連続 3 単語組みからなる素性に変換する。そして、これらの素性を、最大エントロピー法で学習された分類器に入力し、多数の音声操作対象の機能の中から、適切な機能に分類する。なお、図 3 に、認識結果に基づくアクション分類の例を示した。

2.2.3 拡張音声コマンド方式の利点

大語彙連続音声認識では、常に 100% の認識率が得られるわけではない。しかし、後段の

*2 これらの手法間の差異を論ずることは、本報告の目的ではない。

表 1 ドメインごとの機能の数と収集した発話数
Table 1 Target domains and numbers of examples for each domain.

ドメイン	オーディオ	エアコン	ナビ
機能	カーオーディオ操作	空調操作	周辺検索・地図の拡大縮小等
機能の数	51	31	125
収集した発話例の数	9,070	19,605	28,130

アクション分類は、文全体ではなく、認識結果に含まれる単語そのもの、連続 2 単語組み、連続 3 単語組みという局所情報の組み合わせを利用するため、助詞の脱落や挿入などの認識誤りが混入しても、ユーザが意図した機能を正しく起動することが可能となる。このように拡張音声コマンド方式を利用することで、正確にコマンドを覚えていなくても車載機器を操作できる音声入力手法の実現が期待できる。次章以降では、実際にユーザの発話を収集・分析し、拡張音声コマンド方式の可能性を検証する。

3. 車内ユーザ発話例の収集

自動車内でユーザが実際に車載機器操作のためにどのような発話をするか、ということ进行调查するために、252 名の被験者を対象として、ユーザの発話例を収集した。

3.1 対象とした車載機器

音声認識機能の付いたある市販車について、音声で操作可能なオーディオ、エアコン、ナビの 3 分野（以下、「ドメイン」と呼ぶ。）を調査の対象とした。各々のドメインについて、対象とした機能の数は表 1 の通りである。なお、各々の機能に対して、音声コマンド方式で利用する 1 つ以上のコマンドが事前に定義されている。

3.2 状況テンプレートを利用した被験者からのデータ収集

各々の機能を利用する状況を説明したテキスト、絵、およびその組み合わせを準備した。これを、「状況テンプレート」と呼ぶ^{*1}。図 4 に状況テンプレートの例を示した。被験者は状況テンプレートに含まれる表現に影響を受けてしまうことがあるため、表現や絵を変更した複数の状況テンプレートを各々の機能について準備した。

状況テンプレートを利用し、多数の被験者の協力を得て、以下の手順で発話例を収集した。なお、図 5 にデータ収集の流れを示した。

手順 1. 被験者への指示： 被験者に対して、以下の指示を与える。「あなたは車を運転しています。この車の機能は音声認識で操作できます。今から提示される状況になったとき、自分の要求を満たすために、あなたなら何と命令しますか。ただし、相手はただの機械です。コマンドの認識にしか対応していません。また、一文で入力してもらわないと対

*1 上述した先行研究では「状況想定プレート」と呼ばれている¹²⁾。

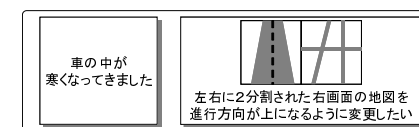


図 4 状況テンプレートの例
(左がテキストのみの例、右が絵とテキストの組み合わせの例)

Fig. 4 Example of situation template.
(The left example shows text only. The right example presents text and figure.)

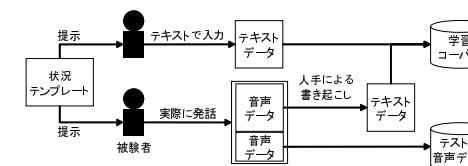


図 5 データ収集の流れ
Fig. 5 Flow of data collection.

応できません。」また、「オーディオオン」「次の曲」というような実際のコマンドを 5 個程度例示する。

手順 2. 状況テンプレートの提示： 各々の機能について、被験者に状況テンプレートを提示し、その状況での発話を考えてもらう。

手順 3. 被験者によるキーボード入力、もしくは発話： 被験者に、状況テンプレートに対応した発話をキーボードで入力してもらおう。また、一部の被験者には、実際に発話してもらい、それを録音する。録音された音声を人手で書き起こし、テキストデータとする。人間はキーボードで入力する場合と実際に発話する場合では、言い回しが異なると考えられるので、実際に発話してもらった音声を収集の方が好ましいが、被験者への負担、および書き起こしのコストが大きいため、全体の約 1% を実発話から収集した。この手順で、機能ごとに 100 から 1,500 程度の発話例を収集した。ドメインごとに収集した発話例の数は、表 1 に示した通りである。なお、図 5 で「学習コーパス」と記述されているデータが、「発話例」に相当する。また、テスト用音声データも、図 5 に示したように、同様の手順で収集した。

4. 車内ユーザ発話の定量的・定性的分析

3 章で収集した発話例・テストデータに基づき、自動車内でのユーザの発話傾向について、定量的・定性的に分析を行った。

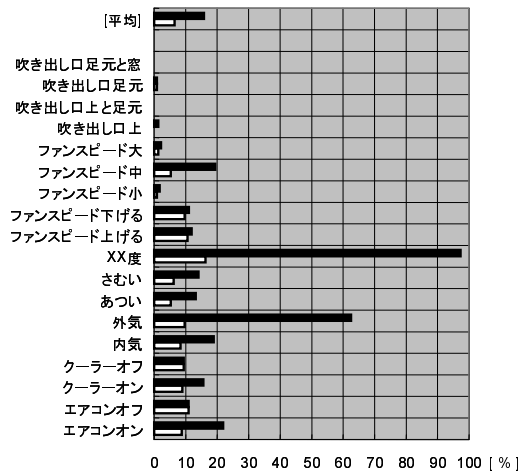


図 6 発話例と音声コマンド方式のコマンドが一致した割合
(白棒はコマンドと完全一致した場合、黒棒は発話例がコマンドを含む場合)

Fig. 6 Ratio that voice commands are exactly matched with and included by utterance examples.
(White bars show exact matches. Black bars show partial matches.)

4.1 定量的分析

音声コマンド方式を利用した場合に、実際のユーザの発話をどの程度受理できるかを定量的に調べた。エアコンドメインについて、各機能に対する発話例が、音声コマンド方式用に用意された文法で受理できるかどうかを調べ、図 6 にこの結果の一部を示した。「コマンドと完全一致」(白棒)は発話例が実際にその機能に割り当てられているコマンドと完全に一致した割合、「コマンドを含む」(黒棒)は発話例が実際のコマンドを含んでいた割合である^{*1}。縦の行が各々の機能に相当し、各機能に対する代表的なコマンドを記述している。1つの機能に対して複数のコマンドが割り当てられている場合もあり、その場合にはすべてのコマンドを正解として取り扱っている。

図 6 を見ると、多くの機能において完全一致は 10%以下にとどまっている。また、コマンドを含む発話例の割合も低い^{*2}。全体の平均は、完全一致が 6.5%、コマンドを含む場合

*1 ある機能に対して、「クーラーオン」というコマンドが割り当てられている場合に、「クーラーオンにして」という発話例は「コマンドを含む」と判断する。

*2 「XX 度」というコマンドが割り当てられている機能については、ユーザの発話例は高い割合でコマンドを含んでいる。この機能は、エアコンの温度設定を行うものであり、多くのユーザは設定する温度を含んだ発話、例えば「XX 度にして」「温度 XX 度」など、を行うので、この機能については高い割合でコマンドを含む結果となった。

が 15.8%であった。これは、ユーザに対して「コマンドにしか対応していない」という指示を与えても、ユーザは直観的には事前に定義されたコマンドを正しく推測することはできず、マニュアルを参照し各コマンドを正しく記憶しなければ、音声コマンド方式を使いこなすことが出来ないことを表している。市販車に用意されているコマンドの総数は数百にも及ぶため、これらを正しく記憶すること自体、困難であろう。これまでメーカー、あるいは開発者が期待していたようには音声入力の実用が広がらなかった原因の一端を、実際のユーザの発話例を定量的に分析することで示すことができた。

4.2 定性的分析

ユーザの発話傾向を定性的に評価するため、表 2 に示す 5 段階のレベル (1~5) と、2 種類の扱うことができないレベル (Ambiguous, Bad) を定義した。そして、各々のレベルに対して、文法に基づく音声コマンド方式で対応できるかどうか、ということを示す 3 列目に、また、4, 5 列目には、拡張音声コマンド方式での大語彙連続音声認識 (認識)、自然言語理解 (理解) が可能かどうかを「○」、「△」、「×」で示した。例えば、音声コマンド方式ではレベル 1 には対応でき、文法を増強することによりレベル 2 にもある程度対応できるが、それ以上のレベルについては対応できないことを示している。拡張音声コマンド方式の場合、レベル 4 以降では、認識対象外の単語が出現する可能性があり、完全に認識できない場合があるため、大語彙連続音声認識は「△」とした。自然言語理解についても、レベル 4 では、認識対象外の単語の影響で誤りが増大する可能性があるため、「△」としている。レベル 5 以降については、理解に高度な背景知識を利用する必要があるなど、入力された一文の発話だけでは正しい判断が出来ない。このように、拡張音声コマンド方式を利用した場合には、レベル 3 までは対応が可能であり、レベル 4 についても一部の発話には対応できる可能性がある。

次に、この分類に基づいて、5.2 節で利用する評価実験用のテストデータ 1,584 発話 (表 4) がどのレベルに属するか、ということ調査した。レベル 1 については、音声コマンド方式用の文法と完全一致する、という評価基準で機械的に調べることができる。その他のレベルについては、人手で分類を行った。図 7 に、各々のドメインについて、各レベルの割合がどの程度であったか、ということ積み上げ棒グラフで示した。これを見ると、すべてのドメインにおいて、レベル 1、つまり文法による音声コマンド方式で対応できる発話の割合は低い。これは図 6 で示した知見と一致する。逆に、ユーザの発話の大部分 (95%以上) はレベル 1 からレベル 4 までに含まれている。特にオーディオ、エアコンドメインについては、レベル 3 までで 96%を超えている。これは、実際のユーザの発話の大部分を拡張音声コマンド方式で処理することができる、ということを示唆している。一方、ナビドメインは

表 2 車内ユーザ発話の分類
Table 2 Classification of user utterances in automobile.

レベル	発話内容	音声 コマンド	拡張音声コマンド		例
			認識	理解	
1	音声コマンド	○	○	○	近くのレストラン
2	事前に想定できる言い回し	△	○	○	近くのレストランを検索
3	事前に準備できる言い回しではないが 意味が十分に伝わる	×	○	○	近隣レストラン検索開始して下さい レストランで食事したい
4	冗長な表現や 認識対象外の未知語を含む	×	△	△	行列のできるレストランでランチを食べたいなあ コンビニまでコーラを買いに行く
5	理解に高度な背景知識を 必要とする発話	×	△	×	いつもの蕎麦屋 あのときの鰻
Ambiguous	他の機能と区別できない	×	△	×	ちょっと上げて*
Bad	内容・意味が不明	×	△	×	これいいね

*「温度」を上げるのか、「音量」を上げるのかを判断することができない。

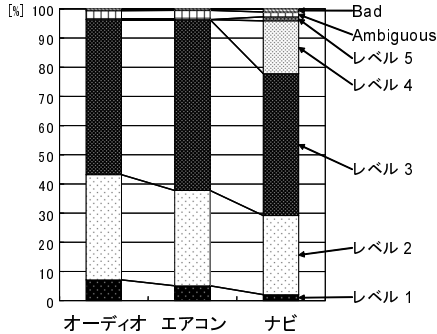


図 7 各ドメインの各レベルの割合
Fig.7 Ratio of utterances in each level by each domain.

他ドメインよりも指示内容が複雑であるためにレベル 4 の割合が増大しており、拡張音声コマンド方式にとっても困難なタスクとなることがわかる。

5. 評価実験

3.2 節で収集したテストデータを利用して、エアコンドメインについて、既存の文法に基づく音声コマンド方式と、拡張音声コマンド方式の性能の比較を行った。さらに、拡張音声コマンド方式を利用して、オーディオ、エアコン、ナビのすべてのドメインについて性能の評価を行った。なお、性能の評価は、ユーザが発話で意図した機能が実行された割合を表す「タスク達成率」を利用して行った。

$$\text{タスク達成率} = 100 \times \text{意図した機能を起動できた発話数} / \text{総発話数} [\%]$$

5.1 音声コマンド方式と拡張音声コマンド方式の比較 (予備実験)

音声コマンド方式と拡張音声コマンド方式を比較するために、エアコンドメインで予備実験を行った。2 種類の方式の実装方法は以下の通りである。

音声コマンド方式： 各機能に割り当てられているコマンドをすべて認識文法として記述する。コマンドと機能が紐付けられており、認識されたコマンドに対応する機能が起動される。

拡張音声コマンド方式： 収集した学習データ (発話例) から、認識用の単語 3-gram モデル^{*1}と、理解用の識別器^{*2}を学習する。これらを利用して、大語彙連続音声認識、自然

表 3 音声コマンド方式と拡張音声コマンド方式を利用した場合のタスク達成率の比較 (エアコンドメイン)
Table 3 Task completion ratio with voice command and expanded voice command.
(Climate control domain.)

	音声コマンド方式	拡張音声コマンド方式
走行雑音なし (21.9dB)	23.0 %	95.0 %
走行雑音あり (4.4dB)	23.5 %	93.0 %

言語理解を行い、ユーザの意図した機能を判断する。

この実験では、雑音がない環境で、10 名の被験者から合計 230 のテストデータを収集した。収集した音声をそのまま利用する場合と、走行雑音を重畳した場合について、2 種類の方式の性能を比較し、その結果を表 3 に示した。重畳雑音の有無にかかわらず、拡張音声コマンド方式を利用した場合のタスク達成率は、音声コマンド方式を利用した場合よりも優れており、拡張音声コマンド方式の優位性を示すことができた。

5.2 拡張音声コマンド方式による 3 ドメインの機能の実装と評価 (実車による評価)

次に拡張音声コマンド方式を利用して、3.1 節で説明した車載機器の 3 ドメインの機能を実装した。各ドメインの機能の数は、表 1 の通りである。テスト音声は、市販車の中で運転席に被験者が座り、実際の車載マイクの位置に録音用マイクを据え付けて収録した。また、音声収録時には、車内エアコンのファンを回し、実際に利用される環境に近い環境を再現した。被験者の数は男性 5 名、女性 2 名であり、収集したテスト音声の総数は 1,584 である。なお、すべてのテスト時に、3 個のドメインの認識用モデルを同時に起動した。

表 4 に各ドメインの大語彙連続音声認識の結果 (文字誤り率・文誤り率) とタスク達成率を示した。この結果を見ると、すべてのドメインにおいて 92%以上のタスク達成率を得る

*1 Modified Kneser-Ney スムージング¹⁶⁾ を利用した。

*2 最大エントロピー法²¹⁾ を利用した。

表 4 拡張音声コマンド方式の市販車での評価実験の結果
Table 4 Result of experiments of expanded voice command with a stock car.

	オーディオ	エアコン	ナビ
テスト発声の数	320	205	1059
認識率			
文字誤り率 [%]	4.3	6.5	7.2
文誤り率 [%]	22.2	28.3	31.3
タスク達成率 [%]	94.4	96.1	92.2

ことができた。特に4章の分析において、拡張音声コマンドにとっても対応が難しい可能性が示されたナビドメインでも92.2%という満足すべき結果が得られた。また、文誤り率に対して、タスクの失敗率は低く抑えられている。例えばナビドメインを見てみると、31.3%の文誤り率に対して、タスク達成率は92.2% (誤り率が7.8%) となっている。これは、拡張音声コマンド方式が音声認識の誤りに対して頑健であることを示している。

6. おわりに

車載機器音声操作のためのユーザの発話傾向に着目し、実際にユーザの発話例を大量に収集した。そして、収集したデータを定量的・定性的に分析し、以下を明らかにした。

- ユーザの発話は多種多様であり、定型句しか受理できない音声コマンド方式では十分に対応することができない。
- ユーザの発話を分類すると、大部分が大語彙連続音声認識と自然言語理解を組み合わせた拡張音声コマンド方式で対応することができる発話である。

さらに、拡張音声コマンド方式に基づく音声入力システムを実際に構築し、評価実験を通じて、「正確にコマンドを覚えなくても利用できる直観的な音声入力手法」を実現できる見通しを得た。

今後、従来の方式に比べて多くの計算機資源を必要とすることや、学習データが必要となるため新規コマンドの追加が容易ではないこと等の問題点を改善し、本方式の適用範囲を広げてゆきたい。

謝辞 日本アイ・ビー・エム株式会社東京基礎研究所音声グループ各位のご協力を深謝いたします。また、データ収集にご協力いただいた関係諸氏に感謝いたします。

参考文献

1) 新田 他: “音声言語インタフェースのための情報処理学会試行標準”, 情報処理学会誌, **47**, 7, pp. 762-767 (2006).
2) 大淵 他: “音声認識技術の実用化に向けた自動車内実環境での評価実験”, 情報処理学会研究報告, 2006-SLP-61-1, pp. 1-6 (2006).

3) L. J. Stifelman *et al.*: “VoiceNotes: a speech interface for a hand-held voice note-taker”, Proc. CHI, pp. 179-186 (1993).
4) R. Sicconi *et al.*: “Honda next generation speech user interface”, Proc. SAE World Congress (2009).
5) 熊井 他: “機能構造と連続キーワード入力を利用した音声インタフェースのユーザビリティ評価”, 日本音響学会秋季講演論文集, No. 2-3-6, pp. 67-70 (2007).
6) P. Brennan *et al.*: “Should we or shouldn't we use spoken commands in voice interfaces?”, Proc. CHI, pp. 369-372 (1991).
7) 西村: “(招待講演) 音声認識ビジネスの現状と将来展望”, 情報処理学会研究報告, 2005-SLP-55-3, pp. 13-15 (2005).
8) 神沼: “自動車用音声インタフェースへの期待”, 情報処理学会研究報告, 2006-SLP-63-9-2, pp. 47-48 (2006).
9) 瀧 他: “カーナビにおける音声インタフェースの評価 - 語彙外発話の状況と対応案に関して -”, 電子情報通信学会技術研究報告, SP2008-152, pp. 57-62 (2008).
10) 渡辺 他: “音節認識を用いたゆ度補正による未知発話のリジェクション”, 電子情報通信学会論文誌, **J75-DII**, 12, pp. 2002-2009 (1992).
11) 北岡 他: “認識結果の正解確率に基づく信頼度とリジェクション”, 電子情報通信学会論文誌, **J83-DII**, 11, pp. 2160-2170 (2000).
12) 河口 他: “CIAIR 実走行車内音声データベース”, 電子情報通信学会技術研究報告, SP2003-136, pp. 139-144 (2003).
13) 李 他: “模擬走行中の車室内における音声インタフェースの感性評価”, ヒューマンインタフェースシンポジウム予稿集, pp. 333-336 (2005).
14) 石川: “UI設計とユーザビリティ~音声インタフェースの課題~”, 情報処理学会研究報告, 2007-SLP-68-7-3 (2007).
15) M. Novak: “Towards large vocabulary ASR on embedded platforms”, Proc. Interspeech, pp. 2309-2312 (2004).
16) S. F. Chen *et al.*: “An empirical study of smoothing techniques for language modeling”, Computer Speech and Language, **13**, 4, pp. 359-393 (1999).
17) H.-K. J. Kuo *et al.*: “Discriminative training of natural language call routers”, IEEE Transactions on Speech and Audio Processing, **11**, 1, pp. 24-35 (2003).
18) 木村 他: “車内音声対話システムのための事例に基づく発話意図推定”, 情報処理学会研究報告, 2002-SLP-40-20, pp. 115-120 (2002).
19) 入江 他: “意図タグつきコーパスを用いた発話意図推定手法”, 人工知能学会研究会資料, SIG-SLUD-A301-2, pp. 7-12 (2003).
20) J. R. Bellegarda *et al.*: “Toward unconstrained command and control: Data-driven semantic inference”, Proc. ICSLP, pp. 1258-1261 (2000).
21) A. L. Berger *et al.*: “A maximum entropy approach to natural language processing”, Proc. Computational Linguistics, **22**, 1, pp. 39-71 (1996).