

状況と文脈を利用した音声対話型車載インタフェースの構築と評価

由浅 裕規[†], 水野 智士[†], 伊藤 敏彦[†], 甲斐 充彦[‡], 小西 達裕[†], 伊東 幸宏[†]

[†] 静岡大学情報学部 〒432-8011 静岡県浜松市城北3丁目5-1

[‡] 静岡大学工学部 〒432-8561 静岡県浜松市城北3丁目5-1

E-mail: [†] {cs8092,cs9082,t-itoh,konishi,itoh}@cs.inf.shizuoka.ac.jp, [‡] kai@sys.eng.shizuoka.ac.jp

あらまし 本稿では、状況及び文脈を利用した言語理解手法をもつ音声対話システムの構築と、その有効性を示すための評価実験について述べる。システムに対する発話方法を「対象を操作」又は「属性は値」という文体に制限することで認識率の向上を図り、また一入力だけでなく連続的に複数の入力を一発話で行えるようにシステムを構築した。タスクとしてはエアコン・カーステレオの操作を想定し、その音声インタフェースという具体的な状況設定において、被験者20名による評価実験を行った。対話の分析を行うことによって、状況及び文脈を利用した言語理解の有効性、制限された発話方法における言語的な特徴や複数設定発話に関するユーザ意識など、様々な知見を得ることができた。

キーワード 音声対話システム, 音声言語理解, 状況と文脈, 発話方法の制限, 複数発話

Construction and Evaluation of Spoken Dialogue Type Car Interface Using a Situation and the Context

Hiroki Yuasa[†], Satoshi Mizuno[†], Toshihiko Itoh[†], Atsuhiko Kai[‡], Tatsuhiko Konishi[†], Yukihiro Itoh[†]

[†] Faculty of Information, Shizuoka University 3-5-1 Johoku, Hamamatsu-shi, Shizuoka, 432-8011 Japan

[‡] Faculty of Engineering, Shizuoka University 3-5-1 Johoku, Hamamatsu-shi, Shizuoka, 432-8561 Japan

E-mail: [†] {cs8092,cs9082,t-itoh,konishi,itoh}@cs.inf.shizuoka.ac.jp, [‡] kai@sys.eng.shizuoka.ac.jp

Abstract This paper deals with the construction of a spoken dialogue system which interprets an input by using the situation/context. The system has restricting input styles to "Operate an object" "An attribute is a value" in order to achieve higher recognition rate. The system further accepts more than one input in an utterance. We have conducted an evaluation experiment by 20 subjects. The experiment involves operating an air-conditioner and a stereo in a car. By analyzing the collected dialogues, the validity of the language interpretation using the situation/context has been confirmed. In addition, the experiment has revealed linguistic features in such restricted utterances and users' recognition of utterance with more than one command.

Keyword Spoken Dialogue System, Spoken Language Understanding, Situation and Context, Restricting Input Styles, More Than One Input in an Utterance.

1. 背景と目的

近年、音声認識技術や言語処理技術、コンピュータ性能の向上により、音声インタフェースや目的指向型音声対話システムが注目されている。音声インタフェース応用は様々な環境・対象に広がっており、車内でも高度な情報システムの音声操作が可能となってきている。運転という主となるタスクがある状況で、音声を用いて様々な操作が行えるということは、リモコンなど手動での操作よりも安全面上適していると考えられる。しかし、実用化を指向したより高度な音声インタフェースの実現においては、いくつかの解決すべき課題がある。その一つとして、音声インタフェースに用いられる音声認識システムの性能向上と信頼性の確

保が挙げられる。音声認識システムにおける認識性能は、発話様式によって大きな影響を受けることが指摘されている[1]。ユーザにとってより使いやすい対話システムを目指し、自由に発声する話し言葉(自由発話)の音声認識に関する研究についても多くの研究報告がなされている[2-3]が、現在の音声認識処理技術では、自由発話では認識率が急激に低下してしまうことや走行ノイズの影響などによって、誤認識を完全に回避することは困難である。誤認識によりシステムはユーザの発話を正しく理解することができないため、ユーザの意図とは異なる動作・応答をすることになる。その結果、正しく理解された場合よりも対話がスムーズに進まなくなり、ユーザに不快感を与えてしまう。

近年ではこの誤認識という問題に対して、システムの信頼性と頑健性を向上させるため、音声認識結果だけでなく、音声認識結果の信頼度を利用する研究も盛んに行われている[4-6]が、それだけでは誤認識に対して十分に対処できているとは言えない。

そこで我々は、誤認識した場合において、信頼度だけでなくその他にも様々な情報を用いることによって、システム自身が誤認識を正しく判断し修正できる頑健な音声対話システムの構築を目指している[7-8]。

また、誤認識を削減するためのアプローチの一つとして、発話方法を制限することも検討している。発話方法を制限することで、自由発話に比べシステム側も発話を予想しやすく、誤認識も起こりにくくなり、最終的にはユーザ満足度も高くなると考えられる。さらには自分の意図をその発話方法で表現しさえすればシステムは発話を受け入れてくれるため、ユーザはシステムに対する発話方法に迷わずにすむという最低限の安心感を得られることができる。また特定の機器の操作だけでなく、他の機器の操作を行う場合でも通用する発話形式を用いる(インタフェースに統一感を持たせる)ことにより、他の新しい機器を操作する場合に改めて操作方法を取得する必要が無く、その点でも効果的であると考えられる。

本稿では、エアコンやカーステレオといった車載器との音声インタフェースを想定し、様々な対象に適用可能な発話方法で音声認識を行い、その結果と認識信頼度、さらに現在の状況や対話の文脈を用いた理解結果を求める手法について述べる。また評価実験を行うことで、音声認識結果(n-best 1位)のみを用いた結果と我々の戦略を用いた結果との比較、制限された発話方法固有の言語的特徴等を示す。また同時に、これまでの一発話で一入力という単体設定発話に加えて、一発話で複数の入力を連続的に行う事が可能な複数設定発話に関しても、ユーザ意識や実対話での利用傾向及びその効果について調査を行った。

2. タスクと発話方法

本研究では車載器との音声インタフェースを想定しているため、タスクは車内のエアコンとカーステレオの操作である。発話方法は「対象物を操作」、「属性は値」という二つの文体に限定した。本システムに対する発話方法例を表1に、本システムで操作可能なエアコンとカーステレオの属性を表2に示す。

表1 発話方法

対象物を操作	エアコンをつける	カーステレオを消す
属性は値	温度は24度	音量はレベル3
属性を操作	温度を上げる	音量を下げる

表2 対象物(種別)と対応する属性

対象物 種別	エアコン	カーステレオ	
	—	ラジオ	CD(MD)
属性	温度	音量	
	風力	ラジオの種類	ディスク名
	風向	放送局	ディスク番号
	フロント デフォッガー	—	曲名
	リア デフォッガー	—	曲番号
	空気循環	—	再生モード

システムに対しては、「温度を上げる」のような単体設定発話だけでなく、「温度を上げる、風力は強」のように一発話で複数の入力を連続的に行う複数設定発話も可能である。また、システムの応答に対して肯定語(「はい」など)や否定語(「いいえ」など)を発話(Yes-No発話)することもでき、これら肯定語や否定語は、設定発話と組み合わせて発話することも可能である。ただし、操作対象を変更する際(カーステレオ⇔エアコン)には、「エアコン(カーステレオ)を〇〇」といった発話が必要である。図1にシステムとの対話例を示す。

U1: エアコンをつける。	[単体設定発話]
S1: エアコンをつけてもよろしいですか?	
U2: はい。	[肯定語]
S2: エアコンをつけました。	
U3: カーステレオをつける、CDをつける。	[複数設定発話]
S3: カーステレオをつけて、CDをつけてもよろしいですか?	
U4: はい、音量を上げる。	[肯定語+設定発話]
S4: カーステレオをつけて、CDをつけました。音量を下げてよろしいですか?	
U5: いいえ。	[否定語]
S5: 音量を上げてよろしいですか?	
U6: うん。	
S6: 音量をレベル3にしました。	
U7: エアコンを操作。	[操作対象変更発話]
S7: エアコンを操作でよろしいですか?	
U8: OKです。	
S8: エアコンを操作にしました。	
U9: 風力は強。	
S9: 風力は中でよろしいですか?	
U10: いいえ、風力は強。	[否定語+設定発話]
S10: 風力は強でよろしいですか?	

図1 システムとの対話例

3. システム構成

図2にシステムの構成図を示す。本研究で構築したシステムは、音声認識部、信頼度生成部、言語理解部、応答生成部、GUI表示部、音声合成部から成る。各モジュール間のデータの受け渡しにはTCP/IPを利用している。ただし、GUI表示部と音声合成部の間のみクリップボードによる受け渡しを行っている。

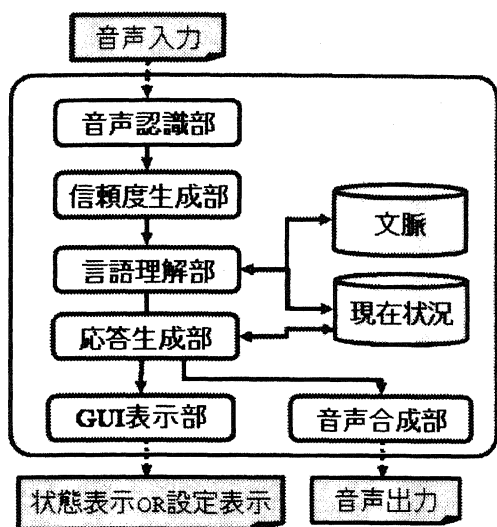


図2 システム構成

3.1. 音声認識部

音声認識部では、入力音声を音声認識し、音響的な尤度で順位付けされた複数候補からなる認識結果 (n-best) を生成する。生成された n-best (n=20) は信頼度生成部へと渡される。本システムでは音声認識器として SPOJUS[9] を用いる。SPOJUS は豊橋技術科学大学の中川研究室で開発された、UNIX または Windows OS 上で動作する日本語連続音声認識システムである。

3.2. 信頼度生成部

信頼度生成部では音声認識部で生成された n-best をもとに、単語/クラスのそれぞれについて信頼度を生成する。この音声認識結果の信頼度は、音響的な尤度と n-best 中の存在率から事後確率に基づく尺度として計算される。この信頼度計算法は、[4] に基づいている。

このように生成された信頼度付きの n-best は言語理解部へと渡される。

3.3. 言語理解部

言語理解部では生成された信頼度付きの n-best に現在状況及び文脈を用いたスコア計算を適用し、スコア付き理解結果 (理解内容) を生成する。言語理解部で生成された理解内容は、応答生成部へ送られ、システム応答に用いられる。理解内容は、(i) クラスパターンの決定、(ii) 決定したクラスパターンに基づく単語理解、の過程を経て生成される。

(i) クラスパターンの決定

言語理解を行う際、まずどのクラスに対する発話が行われたのかを特定する。信頼度付きの n-best は「音量」や「温度」のような『属性クラス』と、「上げる」や「23度」といった『属性の操作(値)クラス』の単語から成る。例えば「温度を上げる」といった発話は『温

度属性』クラスと『温度操作』クラスの組み合わせから成る。この2種類のクラスの組み合わせ(クラスパターン候補)から、理解内容に用いるクラスパターンを特定する。その際、現在どこに対話のフォーカスが当たっているのかといった“話題の対象”にあった、最も順位の高いクラスパターン候補を今回のクラスパターンとする。例えば図3のように、「話題の対象」がエアコンである場合、『音量属性』のクラスパターンは発話されるはずがないため誤認識と判断され、“話題の対象”と合致する『温度属性』のクラスパターンが採用されることになる。

また、n-best 中の全ての単語は認識履歴として、信頼度はスコアとして履歴 DB に保存される。スコアは、これまで発話された可能性を示す数値として用いる。

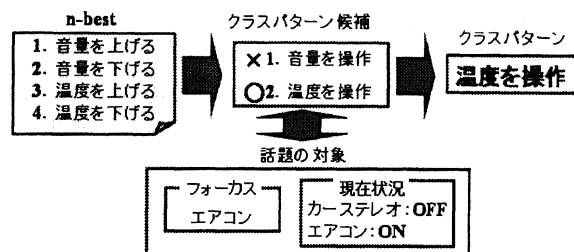


図3 クラスパターンの決定

(ii) 単語理解

理解内容に用いるクラスパターンが決定したら、図4のようにそのクラスの履歴 DB から理解内容に用いる単語を特定する。

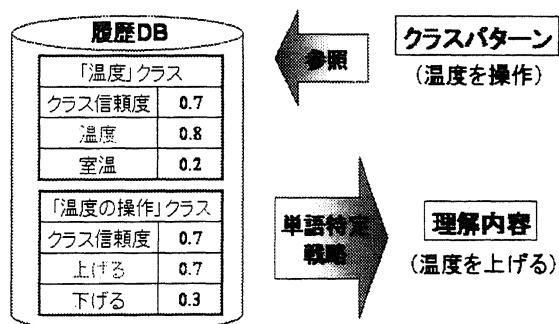


図4 単語理解

その際、主に次の2つの戦略を適用する。

◆戦略1 前回応答単語の不採用

クラスパターンが前回と同様の場合、前回ユーザーに応答した単語はスコアを下げ、その単語を理解内容に用いないようにする。これは、同じ応答の繰り返しを防ぐためである。

◆戦略2 現在状況との矛盾

次に、現在状況と矛盾のある単語もスコアを下げ、理解内容に用いないようにする。現在状況と矛盾していると判断するための適用基準を以下に示す。

- 現在状況値と一致した値単語
- 対象(属性) が ON(OFF)の場合の「つける(消す)」

系操作単語

- 最小値, 最大値をもつ属性が最大(最小)値の場合の「上げる(下げる)」系操作単語

上記の戦略を経てクラス内で最もスコアの高い単語を理解内容として用いる。設定発話の場合、理解内容は属性単語、値(操作)単語それぞれ生成されるのだが、戦略によって信頼できる単語が存在しない(クラス内の全ての単語スコア=0)場合も出てくる。そこで生成された理解内容は、そのスコアで判断された結果(スコア=0なら0, そうでなければ1)とともに応答生成部へ送られ、応答文生成に利用される(表3参照)。肯定語や否定語といった Yes-No 発話の場合は対話に与える影響も大きいので、Yes-No 発話が本当に信頼できるのか、そのスコアとあらかじめ設定された閾値(我々の経験からヒューリスティックに決定)とで判断を行う。スコアが閾値より高い場合は信頼できる(判定結果=1)、低い場合は信頼できない(判定結果=0)となる。

また今回の理解内容が「いいえ」など否定語単独で、かつその発話が信頼できる場合(判定結果=1)、前回のクラスパターンを用いてもう一度理解内容生成を行う。

3.4. 応答生成部

応答生成部では、言語理解部で生成された理解内容とその判定結果から、表3の応答パターンに基づいてシステムの応答文を生成する。生成された応答文はGUI表示部と音声合成部へ送られる。設定完了の場合(肯定語が信頼できる)なら、理解内容も設定内容としてGUI表示部に送られる。その際、理解内容に「操作」系の単語がある場合は、現在状況を利用して「操作」を「値」に変換した結果が応答文及び設定内容として送られる。この変換を行うことで、ユーザに設定後の状況をより分かりやすく伝えることができると考える。

表3 応答パターン

判定結果	応答文
[1,1]	対象物(属性)を操作(値)でよろしいですか?
[1,0]	対象物(属性)をどうしますか?
[0,1]	何を操作(値に)しますか?
[0,0]	もう一度発話して下さい

3.5. GUI表示部

GUI表示部では、応答生成部で作成された応答文、設定内容の表示を行う。GUIは設定画面と状態表示画面の2種類から成る。ユーザがシステムに発話を行うと設定画面へと移行し、設定完了後、一定時間発話が行われないと状態表示画面へと移行する。設定画面では、話題の対象となっている対象物(エアコン or カーステレオ)の属性項目をリスト形式で表示することによって、発話方法とインタフェースとの統一を図って

いる。状態表示画面では、エアコンとカーステレオ両方の現在の設定状況を、図などを用いてユーザに示す。

3.6. 音声合成部

応答生成部によって生成された応答文は、GUI表示部でテキスト表示されると同時に音声合成されユーザに伝えられる。音声合成部には日本 IBM 製の ProTalker97 を使用する。

4. 評価実験

エアコン、カーステレオの設定を想定したオンラインでの評価実験の方法とその結果を示す。実験後の被験者アンケートから主観的評価を、また音声対話データの分析結果から客観的評価を行った。

4.1. 実験方法

本システムは運転中における音声での操作を想定しているため、被験者には運転操作を行ってもらいながらこちらで設定したタスクを実行してもらった。運転操作は安全性の問題からドライブシミュレーション(ゲーム)を使用し、単純なコース(オーバルコース:楕円形コース)を一定速度(100km/h)で走るものとした。ただし、できるだけ実際の運転状況に近づけるため、ハンドル・アクセル・ブレーキを用意し、画面サイズもプロジェクターを使うことにより実サイズに近づけた。このように設定したドライブシミュレーションの操作は、実車の運転操作と同程度の心理的負担を被験者に与えており、その心理的負担度は、実車を60km/hで走行させる場合と同程度であることが既に証明されている[10]。本システムの音声認識器(SPOJUS)は発話の前に毎回音声入力開始ボタンを押す必要があるが、ハンドルに音声入力開始ボタンを取り付けることで、運転に支障がでないよう配慮した。

被験者は情報系大学院)生20名であり、音声対話システムに関する知識はまったくない。

次に実験で用いたタスクについて説明を行う。本システムでは、エアコンに関する操作が7操作、カーステレオに関する操作が13操作で計20操作が可能となっている。この中から12操作を行ってもらうことで1タスク達成とし、図5に示すようなタスク提示文を実験中にこちらからタイミングを見計らって提示した。タスクは4パターン用意し、被験者には複数設定発話も可能なシステム(S1)と、単体設定発話しかできないシステム(S2)でそれぞれ1タスクずつ、計2タスク(24操作)を、ランダムな試行順で行ってもらった。

被験者には実験前に本システムで可能な操作についてよく覚えてもらい、実際に自由に扱ってもらうことでシステム及び発話方法に慣れてもらった。

1. エアコンの操作	7. ラジオの種類
2. 温度の操作	8. 放送局の操作
3. 風力の操作	9. 音量の操作
4. カーステレオの操作	10. エアコンの操作
5. CD の操作	11. 温度の操作
6. ラジオの操作	12. 風向の操作
[タスク提示文 1] なんだか寒いなあ。エアコンつけて、車内の温度を上げようっと。温度を最大の 30 度に設定して、風力も上げよう。(1, 2, 3)	
[タスク提示文 2] 音楽がついてないと寂しいな。カーステレオの CD でも聴こうかな。(4, 5)	
[タスク提示文 3] CD にももう飽きたな。ラジオに変えよう。FM 放送より AM 放送が聴きたいな。(6, 7)	
[タスク提示文 4] この放送局面白くないな。別の放送局に変えよう。ちょっと音量が小さいかな。(8, 9)	
[タスク提示文 5] もう十分暖かくなったな。エアコンの温度を下げよう。でも足元はまだ寒いなあ。(10, 11, 12)	

図 5 タスクシナリオ例

4.2. 実験結果

主観的評価(実験後の被験者アンケート)と客観的評価(言語的特徴の分析)の結果を示す。

4.2.1. 主観的評価

実験後の被験者アンケートから、システムに対する主観的評価を調査した。被験者には質問に対して 7 段階で評価(数値が高いほど評価が良い)を行ってもらった。表 4 に主観的評価の結果を示す。

全ての項目において平均値以上の評価値を得ることができた。特に「発話文体の有効性」、「複数設定発話の有効性」、「複数設定発話ができれば便利か」、「車載器としてこのシステムがあれば便利か」といった項目では高い評価を得ることができた。

4.2.2. 客観的評価

実験によって得られた発話データを言語的特徴に着目し分析した結果を示す。我々の想定した範囲内の発話(想定内発話)に関する言語的特徴を表 5 に、全ユーザ発話に関する言語的特徴を表 6 に示す。表中の「認識→理解」は正しい認識結果から正しい理解内容を導いたもの、「認識→誤理解」は正しい認識結果を得たにもかかわらず誤った理解内容を生成したものの、「誤認識→誤理解」は誤った認識結果を訂正できず誤った理解内容を生成したものの、「誤認識→理解」が誤った認識結果から正しい理解内容を導いたものである。また、表 6 の「複数設定可能ターン」はユーザにとって複数設定発話が可能であったターン数を示し、「複

表 4 主観的評価

質問項目	評価値
話しかけやすさ(次の発話がすぐに浮かんだか)	4.9
聞き取りやすさ	5.6
対話のリズム(発話や応答の速さ)	4.9
システムと対話していて安心感があったか	5.0
車載器としてみた場合の音声認識性能	5.1
車載器としてみた場合の言語理解性能	5.4
応答性能(応答内容は理にかなっていたか)	5.4
車載器としてみた場合の全体の性能	5.0
好感度(このシステムは好きか)	4.9
複数設定発話は有効だったか	5.6
何かを操作する場合に今回の発話文体は有効か	5.7
複数設定発話ができれば便利か	6.1
車載器としてこのシステムがあれば便利か	5.5

数設定発話(2回)」は一発話で二つの設定発話を行ったターン数、「複数設定発話(3回)」は一発話で三つの設定発話を行ったターン数である。「複数設定発話ターン」では、複数設定発話可能ターンにおいてどれだけ複数設定発話を試みられたか、その割合を示している。

表 5 の想定内発話における設定発話に関する言語的特徴をみると、認識率 54.7% に対して理解内容一致率 69.7% と、15% の一致率の上昇がみられる。想定内発話全体でも 8.4% の一致率上昇であり、また「誤認識→理解」では 50.2% と、誤認識した発話の半分以上を理解内容一致へと導いている。S1 と S2 を比べてみると、設定発話数では大きな違いは無いが、Yes-No 発話数では S1 の方が少なく、複数設定発話の影響で「はい」などの肯定語発話が少なくすんでいることが分かる。

表 6 の結果を見ると、一致率上昇は 7.8%、「誤認識→理解」は 37.1% と、想定外発話の影響で想定内発話のみに比べ多少落ちてはいるが、それでも戦略の効果が大きく表れている。複数設定発話が行われた割合は 50.9% と、複数設定発話が可能なたーンの 2 回に 1 回は行われており、その影響で S1 の方が少ないターン数でタスク達成ができています。その他の結果としては、全体的に「あの一」や「えーと」といった「間投詞」の出現頻度が非常に少ないことが分かる。これは[11]で示されている、人間対システムとの対話における間投詞の出現頻度と比べてみても明らかである。

5. まとめと今後の課題

状況及び文脈を利用した言語理解手法をもつ音声対話システムの構築と、実際の運転に近い状況を想定した評価実験を行った。また、制限された発話方法及び発話への影響、ユーザの複数設定発話に対する意識の調査やその効果の分析も同時に行った。

収録した対話音声の分析結果として、単純に音声認

表 5 想定内発話における言語的特徴

評価項目	S1	S2	合計
想定内発話全体	590	687	1277
認識率(n-best1位)	406[68.8]	507[73.8]	913[71.5]
誤認識	123[20.8]	108[15.7]	231[18.1]
認識棄却	61[10.3]	72[10.5]	133[10.4]
理解内容一致	469[79.5]	551[80.2]	1020[79.9]
誤理解	60[10.2]	64[9.3]	124[9.7]
認識→理解	403[99.3]	502[99.0]	905[99.1]
認識→誤理解	3[0.7]	5[1.0]	8[0.9]
誤認識→誤理解	57[46.3]	59[54.6]	116[50.2]
誤認識→理解	66[53.7]	49[45.4]	115[49.8]
設定発話	358	357	715
認識率(n-best1位)	190[53.1]	201[56.3]	391[54.7]
誤認識	123[34.4]	108[30.3]	231[32.3]
認識棄却	45[12.6]	48[13.4]	93[13.0]
理解内容一致	253[70.7]	245[68.6]	498[69.7]
誤理解	60[16.8]	64[17.9]	124[17.3]
Yes-No発話	232	330	562
認識率(n-best1位)	216[93.1]	306[92.7]	522[92.9]
誤認識	0[0]	0[0]	0[0]
認識棄却	16[6.9]	24[7.3]	40[7.1]
理解内容一致	216[93.1]	306[92.7]	522[92.9]
誤理解	0[0]	0[0]	0[0]

表 6 全ユーザ発話における言語的特徴

評価項目	S1	S2	合計
全ユーザ発話	662	777	1439
認識率(n-best1位)	407[61.5]	517[66.5]	924[64.2]
誤認識	164[24.8]	162[20.8]	326[22.7]
認識棄却	91[13.7]	98[12.6]	189[13.1]
理解内容一致	476[71.9]	560[72.1]	1036[72.0]
誤理解	95[14.4]	119[15.3]	214[14.9]
認識→理解	404[99.3]	511[98.8]	915[99.0]
認識→誤理解	3[0.7]	6[1.2]	9[1.0]
誤認識→誤理解	92[56.1]	113[69.8]	205[62.9]
誤認識→理解	72[43.9]	49[30.2]	121[37.1]
複数設定可能ターン	167	-	167
複数設定発話(2回)	76	-	76
複数設定発話(3回)	9	-	9
複数設定発話ターン	85[50.9]	-	85[50.9]
間投詞数	0	3	3
言い淀み・言い直し数	9	12	21
ターン数	568	777	1345

識結果(n-bestの第1位)を用いるより、状況や対話の文脈を利用して導き出したシステムの理解内容の方がユーザ発話との高い一致率を得ることができた。制限された発話方法の特徴としては、自由発話に比べ間投詞が出現しにくいいため認識率の向上が望め、また発話方法自体にも高い評価を得ることができた。複数設定発話に関しては、主観的評価の結果からユーザは一度に複数個の設定ができるインタフェースを望んでおり、また実際の対話においても複数設定発話は頻繁に使われ、ターン数の減少に繋がることが分かった。今後は、今回の実験で明らかになった特徴をさらに詳細に分析し、音声対話システムへの応用を検討する予定である。

文 献

- [1]村上仁一, 嵯峨山茂樹: 自由発話音声認識における音響的および言語的な問題点の検討, 日本音響学会音声研究会資料, SP91-100, pp. 71-78 (1991).
- [2]篠崎隆宏, 斉藤洋平, 堀智織, 古井貞照: 話し言葉音声の認識を目指して, 情報処理学会研究報告, SLP-34, pp. 125-130 (2000).
- [3]河原達也: 話し言葉音声認識の概観, 情報処理学会研究報告, SLP-34, pp. 119-123 (2000).
- [4]駒谷和範, 河原達也: 音声対話システムにおける音声認識結果の信頼度の利用法, 日本音響学会講演論文集, 3-5-2, pp. 73-74 (2000).
- [5]駒谷和範, 河原達也: 音声認識結果の信頼度を用いた頑健な混合主導対話の実現法, 情報処理学会研究報告, SLP-30, pp. 39-44 (2000).
- [6]緒方淳, 有木康雄: 音声認識精度向上のための信頼度尺度の比較, 情報処理学会研究報告, SLP-34, pp. 113-118, (2000).
- [7]水谷誠, 伊藤敏彦, 甲斐充彦, 小西達裕, 伊東幸宏: 音声認識の信頼度と対話履歴を利用した最尤推定型言語理解, 情報処理学会研究報告, SLP-45, pp. 113-118 (2003).
- [8]由浅裕規: 信頼度と文脈を用いた音声対話によるカーナビインタフェースの構築, 静岡大学情報学部卒業論文, (2002).
- [9]中川聖一, 甲斐充彦: 文脈自由文法制御による One Pass 型 HMM 連続音声認識法, 電子通信学会論文誌, Vol. J76-D-II, No. 7, pp. 1337-1345, (1993).
- [10]伊藤敏彦, 甲斐充彦, 岩本善行, 水谷誠, 由浅裕規, 小西達裕, 伊東幸宏: 目的地設定タスクにおける対話状況の違いによる言語・音響的特徴の比較, 情報処理学会論文誌, Vol. 43, No. 7, pp. 2118-2129 (2002).
- [11]岩本善行, 伊藤敏彦, 甲斐充彦, 小西達裕, 伊東幸宏: 同時処理タスクの難易度の変化における言語能力への影響, 情報処理学会研究報告, SLP-41, pp. 61-67 (2002).