

TV 番組の録画再生操作における音声対話機能のユーザビリティ評価

井本 和範, 笹島 宗彦*1, 下森 大志, 上原 龍也

{kazunori.imoto|taishi.shimomori|tatsuya.uehara}@toshiba.co.jp, msasa@ei.sanken.osaka-u.ac.jp

(株)東芝 研究開発センター マルチメディアラボラトリー

あらまし 音声対話システムで問題となる想定外発話の入力を軽減する UI 設計法を提案する。ハードディスクレコーダーに代表されるテレビ番組の録画再生タスクをモチーフに、ユーザが操作に迷いそうな場面で音声入力可能な語彙を強調表示して音声操作を支援するシステムを試作した。ユーザ支援を行わないシステムと比較評価実験を行ったところ、提案法によりユーザビリティが向上することを確認した。

A Prototyping of Voice Interface for Video Recording System and its Usability Test

Kazunori IMOTO, Munehiko SASAJIMA*1, Taishi SHIMOMORI, Tatsuya UEHARA
Multimedia Laboratory, Research and Development Center, Toshiba Corporation

Abstract This paper proposes a new method for designing voice interfaces with less difficulty in finding a valid words for controlling a target device. Applying our method, the valid words are displayed with emphasized color and/or size. Users can see what he/she should say to achieve current goal, for example, change a program, refer with new property, seek for another recorded TV-program, and so on. We also made an experiment by comparing two video recording systems, one with our method and the other one without it. Result of the experiment shows our method improves usability of the voice interfaces.

1. はじめに

いわゆる「デジタル家電」をはじめとして家電製品が近年多機能化している。しかしそのヒューマンインタフェースは決して使いやすくなっていない。例えば 2003 年度、東芝家電ご相談センターには 60 万件以上の相談電話が寄せられたが、その半数以上は取り扱い相談であった[1]。これは多機能化している家電製品を一般消費者が使いこなせていないことのひとつの現れだと考えられる。

多機能の家電を操作するインタフェースとして、音声が目されている[2]。家電製品に対して「話しかけて操作する」という文化が無い現状において、音声インタフェースを家電に搭載し一般ユーザでも利用できるようにするためには想定外発話への対応が必須の課題であり、すでにいくつかの研究が進められている[3, 4]。

想定外発話への対応には受理可能な入力の範囲を広げてあらゆる入力に対応する方針と、ユーザを適切に誘導して想定外の発話をさせないようにする方針の 2 つがある[5, 6, 7]。本研究が対象とするテレビ番組の検索の場合、番組名や出演者名など想定される単語の数が非常に多く、あらゆる入

力に対応することは困難である。

しかし誘導だけで適切なインタフェースを構築することも困難である。誘導や確認のための単純なやり取りを増やすことで対話は頑健になるが、インタフェースの操作性は一般に悪くなる。

そこで、多機能家電製品の操作性を音声インタフェースの適用によって向上させるためには、単純なやりとりを必要以上に増やすことなくユーザを暗黙的に誘導する技術が必要である。筆者らは TV 番組の録画再生操作をモチーフとして、操作のために本質的な語彙を必要なタイミングで画面に強調表示しユーザを誘導する方式について検討してきた。また、提案手法に基づいた試作と評価実験を行い、提案方式の有効性を確認した。本稿では音声インタフェースを家電製品の操作に適用する場合の想定外発話の問題を解決するための一方式を提案し、その適用システム評価実験結果について述べる。

2. TV 向け音声インタフェースの課題

2.1. 基本システムの構築

複雑化する家電製品のインタフェースとして音声期待されているが、現状では家電製品に対して「話しかける」行為は一般的ではない。一部の音声認識機能を利用した自動電話応答システムで

*1 現在、大阪大学産業科学研究所 知識システム研究分野

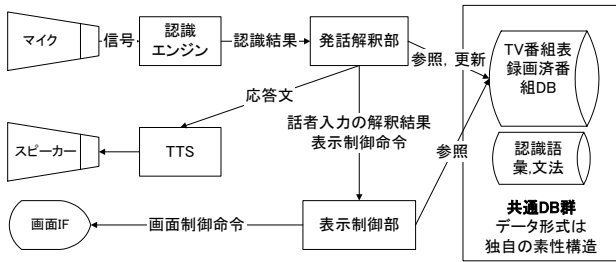


図 1：音声対話による録画再生機の基本構成

```

Program_of(int:ID,          /* 番組データID */
string:ProgramTitle,      /* 番組名 */
string:ProgramSubTitle,  /* 番組副題 */
int:Date,                 /* 日付 */
int:DayOfWeek,           /* 曜日 */
int:StartTime,           /* 開始時刻 */
int:EndTime,             /* 終了時刻 */
String:Channel,          /* チャンネル名 */
String:Genre,            /* ジャンル */
String:SubGenre,         /* サブジャンル */
StringList:ActorList,    /* 出演者リスト */
String:Content,          /* 番組内容 */
... (以下、内部モジュール参照用のフラグなど)).

```

(a)

```

Program_of(4,あすの料理,ハンバーグと肉だんごの甘酢あんかけ,
1016,水,930,1000,HHK,趣味・暮らし,料理,[山田順子],.....1).

```

(b)

図 2：番組データ構造(a)とインスタンス例(b)

は人間が機械に対して音声対話をしているが、この場合のヒューマンインタフェースは受話器であり、そもそも話しかけるために設計された機械であるため我々は違和感を持たない。

家電の音声操作に対してユーザがどのような反応を示すか観察するために、はじめに図 1 のような構成のシステムを試作した。音声認識エンジンには複数単語スポッティング機能を持つものを採用している。試作システムは番組を検索するための発話と装置の状態を制御するための発話との大きく分けて 2 種類の発話を受理する。テレビ番組を検索するための発話には番組名やジャンル名称のように 1 単語から成るものと「金曜日午前 10 時」のように複数の単語から成るものがある。

また対話戦略の原則としてユーザ主導対話を採用する。不慣れた音声インタフェースを利用するユーザを支援する技術としてはシステム主導対話が挙げられる。ユーザに適切な発話を促したり選択肢そのものを提示したりすることで、認識率の向上や対話の流れを破綻させないなどの利点がある。しかしテレビ操作においてユーザを適切に誘導することは困難である。たとえば番組を選択する操作だけを考えても、番組名のほかにチャンネ

ル名、出演者、日時、ジャンル、放送内容など入力が多岐にわたり、事前の予測は困難である。これに対してユーザモデルの利用による絞込みは有効であるが、テレビはお茶の間で複数人が利用する家電であり、複数のユーザモデルを管理することは困難である。

検索対象となる番組データの構造を図 2(a)に示す。これを図 2(b)のように具体化したものがデータベースに蓄積されており、ユーザは各属性について意図する条件を音声入力することによって番組データを検索できる。検索機能として「金曜日」「午前 10 時」のように直前の発話入力による検索結果を次の発話でさらに絞り込む機能と、「なんとか太郎」のように人物名の条件が曖昧でも検索ができるあいまい検索機能を実装した。これら検索機能を利用して、本システムは(1)テレビ番組表からの番組検索と録画(2)録画済番組データベースからの番組検索と再生(3)オンタイムでのテレビの視聴、以上 3 つの機能を実現している。

また、認識語彙については「番組表」「再生モード」など命令語のように不変のものと、番組名や出演者名などのように日によって変化するものがある。前者についてはあらかじめ用意できるが、後者についてはデータベースに蓄積された番組データをもとに、本システムに組み込みの TTS エンジン内部モジュールである「漢字仮名変換エンジン」を利用して半自動で生成する。

2.2. 基本システムによる予備実験

図 1 の構成で試作機を PC 上に実装し、一般から募集した 10 名 (20 代男性 5 名, 20 代女性 5 名) の被験者を対象に、グループインタビューによる評価実験を行った。評価実験機には 8 日分の番組データ (現在、番組情報提供ホームページで標準的に提供されている番組表のサイズ) を搭載した。具体的には、放送局 10 局分の番組データをもとに、図 2 のデータ構造をもつレコードからなるデータベースを作成した。また、認識語彙についてもそれをもとに作成した。番組名や出演者名については 2.1 節で述べた手法で半自動生成した。認識語彙は合わせて約 5000 語程度である。

インタビューは男女別のグループを対象にそれぞれ約 2 時間ずつ実施し、発言を記録して分析した。前半の 1 時間で音声操作に対して被験者が持つイメージの分析と概念レベルでの説明、後半の 1 時間で試作機の説明とデモを行った。

その結果、筆者ら開発者によるデモを見た被験

者は、おおむね「使ってみたい」「早い」「機械操作が苦手な人には便利だ」などの好意的な反応を示した。一方で「初心者向きではない」「使う前に構えて（考えて）しまう」などの意見もあり、試作機のレベルではまだ初心者の捜査に対して十分な支援ができていないことが明らかになった。事実、インタビュー中に10分程度のデモ使用の後「興味のある方は使ってみてください」と被験者による試用を促したが、実際に使ったのは1名だけで、それも出演者名1発話だけの入力であった。

さらにインタビューを通じて、テレビの視聴形態には目的をはっきり決めて見る場合と特に目的無く見る場合の2つがあることと、本試作機は後者に対応していないという問題点が明らかになった。目的すなわち見たい番組を決めて本システムを利用するユーザは、検索条件である番組名や出演者名などをはっきり言える。他方、特に見たい番組が無いままにテレビのスイッチを入れて、何か他のことをしながら環境音楽のようにテレビを利用するユーザも存在する。後者に対しては漠然とした目的を具体化させるための支援策が必要である。例えば番組をスポーツや音楽などジャンルのレベルで選択したいユーザに対しては、今それが発話入力可能か、他にどのようなものがあるかなどを提示することで支援できると考えられる。

3. TV 向け音声インタフェースの改良

3.1. 想定外発話の問題

予備実験が示唆するように音声対話システムが抱える課題として想定外発話への対応がある。想定外の発話は（1）辞書に登録された単語をユーザが知らない（2）辞書にユーザの発する単語を全て網羅できない、の2つが原因となって生じる。TV 番組検索タスクは認識語彙が膨大であるうえ日々内容が更新されるため、この問題がより顕在化する。さらに複数の機能を扱う本タスクでは、対話状態によって受理する語彙が切り替わる。そのため（3）対話状態が異なるため受理できないことも想定外発話の原因となる。

想定外発話への対応には受理可能な入力の範囲を広げてあらゆる入力に対応する方針と、ユーザを適切に誘導して想定外の発話をさせないようにする方針の2つがある。受理可能な入力の範囲を広げるには、詳細なタスク分析によって入力される語彙を網羅的に登録する必要がある。だが全ての語彙を洗い出すことは困難で、語彙の増大が

表 1：対話システムの語彙分類

・コマンド語彙	
・語彙知識を共有	なし
・語彙体系が異なる	全て 例) これを録画
・コンテンツ語彙	
・語彙全体を共有	チャンネル, 日時
・語彙体系が異なる	ジャンル
・語彙集合が異なる	
・複数読み	番組名, 出演者

認識率低下につながるため限界がある。一方、ユーザに想定外の発話をさせないように、対話の主導権を切り替えて誘導する手法は、ユーザの目的と合致した誘導でなければ、冗長な対話を強いる危険がある。そこで本研究では画面に認識語彙を提示することで暗黙的にユーザを誘導する手法を採用する。画面を介して効率よく語彙知識を共有するには、必要十分な語彙をいつでも参照できるようにわかりやすく表示することが重要である。受理できる単語が多く、語彙の表示に画面の多くを割けない本タスクでは、単語の性質に応じて提示する語彙を絞り込む必要がある。また画面に表示される他の文字情報と容易に識別できるように表示方法を工夫する必要がある。

3.2. 認識語彙の分類

本節では認識語彙の性質を整理して、ユーザに提示すべき語彙について考察する。表 1 に認識語彙を分類したものを示す。分類の基準としては 1) コマンド語彙なのかコンテンツ語彙なのか、2) ユーザとシステムに語彙知識の差異はあるか、3) 単語が複数の読みを持つ可能性があるか、を用いた。以下各基準について詳細に説明する。

・コマンド語彙／コンテンツ語彙

コマンド語彙は対話システムの操作に必要なシステム依存の単語群である。使用頻度が高く不変のため、直感に合った、覚えやすい単語を登録することが重要である。コンテンツ語彙は対象タスク依存の単語群で、ジャンル、番組名、出演者など番組の属性情報がこれに当たる。本タスクではコンテンツ語彙の更新頻度が高い。

・ユーザとシステムの語彙知識の差異

想定外発話を軽減するためには、ユーザとシステムの間で認識語彙を共有する必要がある。コマンド語彙では、既存のビデオ機器の操作から類推できる単語を登録することで多くを共有できるが、完全には一致しない。コンテンツ語彙は属性の性質によって傾向が異なる。日時やチャンネルは固定

3件 HITしました	テレビに戻る	番組表
① 10/16	09:30-10:00	料理 あすの料理
② 10/16	11:45-11:50	料理 世界食紀行
③ 10/16	17:15-17:30	料理 夕食クッキング
前のページ これを録画 次のページ		

図3：認識語彙の強調表示の画面例

的で入力される単語も予測できるため共有できる。ジャンルは単語自体をある程度共有できるが、個人によってサブジャンルを含めた語彙体系が異なる。番組名や出演者は、システムが取得した全番組名をユーザは知らず、逆にユーザが記憶している過去の番組をシステムは知らない。これらの属性は語彙知識の集合が異なる。

・単語の読みの多様性

複数の読みを持つ単語に関しては、読みのレベルでも語彙知識の共有が必要である。複数の読みには、基本となる読みの部分文字列で表現される省略形から、愛称など基本の読みと対応が難しいものまでその変動パターンは多岐にわたる。変動パターンによっては基本の読みから自動生成することが難しく、全ての読みをカバーできない可能性もある。本タスクではコンテンツ語彙の番組名や出演者名がこれに当たる。

以上の考察より、使用頻度の高いコマンド語彙やユーザとシステムの間で誤解が生じやすい体系の異なる語彙を優先して表示する必要がある。

3.3. 認識語彙の強調表示

限られた領域になるべく多くの認識語彙を提示するには、表示方法を工夫する必要がある。本システムでは、画面を介してユーザとシステムで番組情報の共有を行うが、その中には認識語彙も含まれている。すでに表示された文字情報を語彙の共有にも利用すれば、視認性の確保にも有効である。そこで前節で重要と判断した単語を認識語彙提示用の領域に表示して、それ以外は、タスク遂行上提示される文字を強調表示することで視認性を確保する。また認識単語は全て同一フォントを用いるなど統一することで、他の文字情報と容易に識別できる。具体的には以下の表示戦略を取る。

- ・受理可能なコマンド語彙は常に強調表示する
- ・語彙体系の異なるコンテンツ語彙は、体系の

表2：TV番組の録画・再生・視聴に関する課題一覧

テレビのチャンネルを切り替えて下さい	コマンド
番組表を表示して下さい	コマンド
金曜日の番組表に切り替えて下さい	コマンド
<番組名>を録画して下さい	指定番組
<番組名>を一週間全部録画して下さい	指定番組
<出演者>の出る番組を録画して下さい	指定番組
<出演者>の出た番組を再生して下さい	指定番組
再生を停止して下さい	コマンド
好きなジャンルの番組を再生して下さい	自由番組
見たい番組を何でも1つ録画して下さい	自由番組

一致する上位概念を強調表示する

- ・その他のコンテンツ語彙は、視認性を考慮してタスク遂行の上で画面表示される単語のみ強調表示する
- ・複数の読みを持つ単語は、読みとの対応が可能な部分文字列のみを強調表示する
- ・番組の絞込み検索時には、絞込みに有効なコンテンツ語彙のみを強調表示する

以上の規則に基づき、番組の絞込み検索において認識語彙が強調された画面の一例を図3に示す(図では下線の単語が認識語彙を表す)。

4. 評価実験

提案手法がユーザビリティをどの程度改善させるか評価するための実験を行った。

4.1. 評価基準

まず音声対話システムの使いやすさを評価する基準について考察する。音声対話システムの使いやすさを評価する指標として、課題を達成するまでの対話時間、発話回数、ターン数などの客観値か、主観的な満足度がよく利用される。従来はこれらの評価基準を既存システムと比較することでユーザビリティ評価を行うのが一般的であった。

一方、複数の機能を実現する本タスクでは、対話状態が複雑になるため、従来以上にシステムとユーザの持つ知識の差が広がる。これを埋めるためには、システムに問い合わせるもしくはマニュアルを参照する(以降ではまとめて質問と呼ぶ)が必要があり、ユーザにとって大きな負担となる。インタフェースの操作性を向上させるには、質問回数を減らすことが必要不可欠と言える。

この質問回数は、例えば対話時間のような既存の評価指標に間接的に影響を与えるものの、ユーザビリティとの関係を調査した例はこれまでにない。そこで我々は機器の使いやすさ、すなわちユーザビリティと質問回数の関係を調査し、また提

表 3：両システムの評価結果

	基本システム	提案システム
(平均)満足度 (1-3)	1.9	2.2
(平均)質問回数	18.1	9.6
(平均)発話数	59.0	50.5

案手法によって質問内容がどのように変化するかについて調査する。

4.2. 実験手順

本節では、実験の手順について詳細に説明する。実験には提案手法を実装したシステム（以降では提案システムと呼ぶ）と実装していない基本システムの2種類を準備し、40人の被験者を2グループに分けて、それぞれ20名ずつで評価を行った。被験者にはまずシステムの概要および簡単な操作方法が書かれたマニュアル(A4 サイズ 3 枚)を手渡した。これにはシステムで実現可能な機能、マイクの操作方法、操作画面のスナップショット、音声コマンドなどが記載されている。なおこの時点では被験者に課題を提示せず、課題の途中で制限なくマニュアルを参照しても構わないことを伝えているため、特定の操作方法のみを暗記する被験者はいなかった。マニュアル参照後に、被験者に表 2 に示す10の課題を提示して上から順にこなしてもらった。コマンド語彙およびコンテンツ語彙を組合せた様々な発話パターンが含まれる課題を設定した。また番組を実際に録画再生する場面を想定して、対象の番組が明確である指定番組検索と、目的が漠然とした自由番組検索の課題を与えた。なお表2で”<>”の表記は、実験時には具体的な固有名詞が与えられていたことを示す。

各課題は途中で放棄せず最後まで実施してもらい、課題の達成を確認してから次の課題に移った。課題ごとに質問回数、質問内容、発話回数、発話内容、認識の成否を記録した。なお実験ではシステムへの問い合わせは、実験に同伴した開発者への質問に置き換えたが、「質問された内容にのみ答える」ように取り決め、必要以上の情報が被験者に伝わらない工夫を施した。全ての課題を終了した時点で、被験者にシステムの使いやすさに対する満足度とその理由を自由回答形式で確認した。

4.3. 実験結果

全ての課題を達成するために発声した発話数、質問回数およびシステムの使いやすさに対する満足度の平均値を表 3 に示す。満足度は1から3までの小数点を含む任意の数を取り、数字が大きいほど評価の高いことを表す。表より全ての項目に

表 4：質問内容の詳細

質問内容	基本システム	提案システム
コマンド	11.5 (63.3%)	5.7 (58.9%)
コンテンツ	3.6 (20.2%)	1.6 (16.8%)
状態	1.7 (9.7%)	1.8 (18.3%)
操作	0.1 (0.6%)	0.2 (1.5%)
タスク	0.8 (4.1%)	0.3 (3.5%)
その他	0.4 (1.9%)	0.1 (1.0%)
総質問数	18.1	9.6

表 5：発話内容の詳細

発話内容	基本システム	提案システム
コマンド系		
成功	28.1 (47.6%)	28.1 (55.6%)
誤認識	2.8 (4.8%)	2.8 (5.5%)
辞書文法外	4.7 (7.9%)	2.8 (5.6%)
非受理	3.0 (5.0%)	1.8 (3.5%)
コンテンツ系		
成功	15.5 (26.2%)	10.4 (20.5%)
誤認識	2.2 (3.6%)	3.3 (6.6%)
辞書文法外	2.6 (4.4%)	0.9 (1.7%)
非受理	0.3 (0.5%)	0.5 (1.0%)
総発話数	59.0	50.5

において提案システムが基本システムを上回る結果となった。各項目に対して有意水準5%のt検定を行ったところ、質問回数および満足度において有意な差が得られた。実験結果を詳細に分析するために、質問を内容ごとに分類したものを表 4 に示す。表中の「コマンド」はコマンド語彙に関する質問、「コンテンツ」はコンテンツ語彙や番組の絞込方法に関する質問、「状態」は、誤認識や誤操作により分からなくなった対話状態を確認する質問を示す。また「操作」は音声入力操作の質問、「タスク」は課題内容を確認する質問に対応している。

基本システムにおける質問回数の分布より「何を言えばよいのか」が大きな問題であることが分かる。提案システムでは、「コマンド」および「コンテンツ」に関する質問は半減している。「コマンド」「コンテンツ」に関する質問が総質問数に占める割合を、有意水準5%のt検定で評価したところ、基本システムと提案システムには有意な差が得られた。これは提案システムが「何を言えばよいのか」の解決に有効であることを示す。

続いて発話内容を分類したものを表 5 に示す。各発話を認識結果に応じて「成功」「誤認識」「辞書文法外」「非受理」に分類して、コマンド語彙とコンテンツ語彙それぞれで集計した。「成功」はシステムが想定した単語が発声されかつ正しく認

表 6：満足度と各評価指標の関係

	認識率	発話回数	質問回数
相関値	0.37	-0.29	-0.42

表 7：満足度の決定に影響を与えた項目(自由回答)

回答項目	発言者数
何を言えばよいか分かりやすい	25
現在の状態が分かりやすい	25
認識性能が高い	16
少ない手順で操作できる	15
操作が覚えやすい	14
悩んだときに誘導してくれる	12
様々な言い方で入力できる	9
過不足ない情報が提供される	9

識できた発話を、「誤認識」は想定した単語であるが、認識エンジンが正しく認識できなかった発話に対応する。また「辞書文法外」は場面には合うが認識辞書もしくは文法に登録されていない発話が、「非受理」は認識辞書には登録されているが、対話状態が異なるために受理できない発話に対応する。表より基本システムの認識失敗率は 26.2% であるが、その中で想定内発話を誤認識した割合は 1/3 程度であるため、想定外発声がかかなり含まれていることがわかる。基本システムと提案システムを比較すると、辞書文法外の発話数および割合はおよそ半減しており、結果として認識率が向上している。辞書文法外発話が総発話数に占める割合を有意水準 5% の t 検定で評価したところ、両システムには有意な差が得られており、発話内容からも提案システムが有効といえる。

4.4. 考察

本実験より認識語彙の強調表示が認識率や質問回数の改善に貢献し、システムのユーザビリティを向上させることを示した。これより認識率や質問回数には満足度との相関があると言えるが、その強さについては分からない。そこで満足度と各評価指標の相関の強さを調べるため、相関分析を行った。表 6 に認識率、発話回数、質問回数と満足度の相関値を示す。

表 6 よりどの評価指標も相関値に大差はなく、質問回数との相関が相対的に高い結果となった。被験者に対して行ったアンケート（満足度を定める要因を自由回答形式で答えてください）から、原因を分析した。表 7 に回答項目と発言者数の関係を示す。表には「何を言えばよいか分かりやすい」「現在の状態が分かりやすい」「悩んだときに誘導してくれる」など質問回数に影響する項目

が挙げられている。これが質問回数と満足度の相関が高い要因と考えられる。しかしこれらは発話回数や認識率にも間接的に影響を与えるため、結果として各評価指標に大差がない結果となった。

5. まとめ

本研究では音声対話システムが抱える想定外発話の問題を取り上げ、TV番組の録画再生操作をモチーフとした対話システムを構築してユーザビリティ評価を行った。語彙の性質に応じて選択的に認識語彙を強調表示する方法を提案し、評価指標として質問回数を導入して、基本システムとの比較実験を行った。その結果、提案手法により満足度および質問回数が改善され、想定外発話が減少することを確認した。また評価指標である質問回数が満足度と相関を持つことを確認した。

しかし実用を考えると、想定外発話をさらに減らす必要がある。提案手法では暗黙的な誘導を行うため、強調表示に気付かず辞書文法外の発話を繰り返すユーザも存在した。多くのユーザに対応するには、音声による誘導などその他の手法を組合せた対策も必要となる。ただその際には多様なユーザの目的に合わせて誘導を切り替え、無駄なやり取りを増やさない工夫が必要である。また対話状態の誤解は非受理発話を誘発するため、現状態を積極的に通知する工夫も必要である。

今後は他のタスクにおいても評価を行い、提案手法の有効性を検証する予定である。

<参考文献>

- [1] http://www.toshiba.co.jp/csr/jp/cs/service_j.htm
- [2] 渡辺裕太, 関口芳廣, 鈴木良弥, “ビデオ装置を例とした家電品の音声対話機能について”, 情処学論, vol. 44, no. 11, pp. 2690-2698, 2003.
- [3] 井上剛, 西崎誠, 小沼知浩, 桑野裕康, 脇田由実, “EPG 情報番組検索向け音声インターフェースにおける認識語彙選択手法の検討” 情報処理学会第 65 回全国大会
- [4] 内田尚和, 常盤大樹, 西末衣, 高木朗, 麻生英樹, 橋本政朋, 森彰, 中島秀之, 伊東幸宏, 小林一郎, 八名和夫, “情報家電の操作のための対話インタフェースの開発”, 情報処理学会 SIG-SLP, vol. 54, pp. 265-270, 2004.
- [5] 山本幹雄, 高木三功, 中川聖一, “メニューによりガイドされた文節単位発声による音声対話システム”, 情処学論, vol. 37, no. 4, pp. 461-470, 1996.
- [6] 桐山伸也, 広瀬啓吉, 峯松信明, “話題知識を導入した文献検索音声対話システム”, 信学論, vol. J85-D-II, no. 5 pp. 863-876, 2002.
- [7] 駒谷和範, 鹿島博晶, 田中克明, 河原龍也, “複合的言語制約に基づくキーフレーズ検出を用いた汎用的なデータベース検索音声対話プラットフォーム”, 情処学論, vol. 44, no. 5, pp. 1333-1342, 2003.