

1カ22-4

早稲田大学大学院理工学研究科

# 博士論文概要

## 論文題目

音声対話インタフェースにおける  
コミュニケーション調整機構の研究

Study on Regulating Communication  
in Spoken Dialogue Interface

申請者

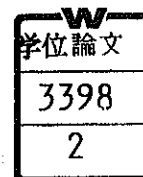
菊池 英明

Hideaki KIKUCHI

情報科学

ヒューマン  
インタフェース

2002年4月



現代社会においては、情報化の進展に伴い様々なレベルでシステムの複雑化が進み、システムを制御する側の人間にとっては扱いが難しくなる一方になっている。特にシステムの利用において特別な知識を要求することの難しい公共的なシステムについては、より汎用的で使い易いユーザインタフェースの必要性が高いといえる。音声による対話は、人間が普段用いるコミュニケーション手段であることから、ユーザインタフェースに適用されることが期待されている。

音声による対話を工学的に実現しようとする研究は、音声の認識や合成、自然言語処理などの要素技術が現実化してきた1970年代頃から多くの研究機関によって行われて来たが、現在なお数多くの課題が残されている。そのうちの一つの大きな問題として、相互作用性の問題があげられよう。音声対話の工学的な実現を目指した研究の多くは、音声認識・音声合成・自然言語処理などの要素技術の応用という観点で始められており、その結果、ユーザによる音声入力からシステムによる音声出力（ないしは行為遂行）という逐次的な処理を想定し目標としていた。しかし近年、要素技術の発展とともにコミュニケーション手段としての音声対話が注目されるようになり、音声対話メディアの利点である相互作用性に焦点をあてて、ユーザである人間と機械とのコミュニケーションをより豊富なものにしようとする研究が盛んになっている。音声対話インタフェースの相互作用性を向上するためのアプローチはさまざまであるが、中でも対話を円滑に進行するためのコミュニケーション調整行為の実現は、最も重要なものの一つといえる。本研究は、コミュニケーション調整行為を実現するためのアーキテクチャの一般的なモデルを構築しようとするものとして位置付けられる。本論文の第1章ではこうした研究の背景と目的を述べる。

第2章では、音声対話インタフェースにおけるコミュニケーションの調整機構において最も重要なものの一つである『発話権』の管理モデルを提案する。発話権とは、対話における発話の順番についてのある種の秩序であり、『発話の番』とも呼ばれる。対話参加者は対話の中で発話権の所在が誰にあるかを意識・無意識に関わらず判断して発話のコントロールを行っているといえる。本研究ではまず、これまで主に言語学の分野で理論付けされてきたこの発話権に対して工学的なモデルを与えて、実際の音声対話システムの上に実現した。具体的には、人間同士の対話音声を集めて分析した結果から、発話権移行の合図となる現象を発見し、それにしたがって、ユーザの発話行為とシステム自身の行為から発話権の所在を推測したうえで状況に応じて発話権を移行させる発話権管理方式を提案した。さらに、発話タイミングの自由を保証した場合に生じる割り込み発話に対して、システムの発話計画とユーザの受理した発話との間に差異が生じないように、発話計画の単位を情報単位とし、発話権の所在を管理することによってどの情報が受聴されたかを常に把握するモデルを提案した。模擬的な対話実験の結果、提案した方式によって、システムが計画した発話とユーザが受聴した発話の差異をなく

しながら割り込みに対処できることが認められた。また、割り込みに対処することにより、ユーザのタスク完了までの所要時間は減少し、積極的な話題提起が増えるなど、音声対話インタフェースとしての利便性が向上することが確認された。この結果から、コミュニケーション調整機構の基本となる発話権管理モデルの有効性と、音声対話インタフェースにおけるコミュニケーション調整の効果を確認できたといえる。

このようにユーザに発話タイミングの自由を保証することによって、ユーザは思い浮かぶ要求を即座に行う行為として遂行できるようになるが、その一方で、ユーザの行為に制約を与えなければ、その結果として問題解決が阻害されたり発話が重複してしまったりなど、対話の進行が滞るケースが増える。このような場合に、例えば問題解決の状況を説明したり、再発話を要求したりするなどの、コミュニケーションの調整をはかるメタ的な発話の生成が対話の円滑な進行に役立つ。本論文の第3章では、コミュニケーション調整機構に必要な機能として、システムや対話の状態を的確なタイミングで提示する状態提示行為制御モデルを提案する。具体的には、システム内部の音声認識や情報検索などの処理精度や処理時間から決定される内部状態の曖昧さ指標の変化に応じて、確認発話をはじめ自己調整発話、状態提示発話などシステム内部の状態を提示する発話を選択し的確なタイミングで出力する。これによって、常に一定のタイミングで状態提示を行うかあるいは全く行わない従来の対話制御方法よりも、ユーザからみたシステムの状態の曖昧さを効率よく低減し、対話の効率や自然性の評価向上につながることを実験により確認した。

状態提示行為制御モデルは、コミュニケーション調整機構において、調整行為の一つである状態提示行為の制御を実現するものである。したがって、コミュニケーション調整機構の一般化においては、調整行為全般についての制御を可能にする枠組みが必要となる。本研究ではこの問題に対して、まず汎用的なプラットフォームとして音声対話インタフェース全体を表現し、システムの内部状態および発話権の状態を次元の一つとする多次元心的状態をそこに導入したうえで、コミュニケーション調整機構の一般化をはかる。

まず第4章に、音声対話制御機構をアーキテクチャレベルで汎用化するための音声対話インタフェース汎用プラットフォームについて述べる。音声対話インタフェース汎用プラットフォームは、音声を入出力手段として問題解決を支援・代行する音声対話インタフェースの開発を容易にするものであり、近年類似の研究が報告されている。本研究では、実現する対話制御の多様性と対話制御規則記述の容易性のトレードオフを考慮して、具体的な行為レベルの上に抽象的なプランレベルを設けて制御の単位とする。これによって、従来法よりも少ない記述量で多様な対話制御を実現可能にする。具体的には、行為レベルの制御規則は有限状態オートマトンで、プランレベルの制御規則は後ろ向き推論規則で表現する。全

ての規則をシステム開発者が記述してもよいが、行為レベルの制御規則をパターンとしてあらかじめ用意し修正可能にしておくことによって、開発者による記述量を削減することを実現した。また、このプラットフォームを用いて、学術論文検索タスク、チケット予約タスクなど異なるタスクについて複数の音声対話システムを実際に構築し、多様な対話制御が実現できることを確認した。

第5章では、第4章に示した状態提示行為制御モデルを、コミュニケーション調整機構として一般化する。状態提示行為をコミュニケーションの調整をはかるメタ的発話行為として一般化する場合、単一チャンネルにおける時間リソース分配の問題がある。つまり、音声対話メディアにおいては発話によって行為を遂行する必要があり、発話権の所在が固定されている（あらかじめ時間リソースが対話参加者に分配されている）対話様式であれば、プランニングによって決定されたメタ発話を通常のやりとりの中に埋め込むことが容易である。しかしながら、発話権の移動が自由な様式の場合には、発話権の所在や内外の状況の変化を予測しなければ通常のやりとりを阻害しないようにメタ発話を出力することが難しい。本研究では、音声対話における状況変化の予測に応じてプランニングにより状況に適したメタ発話生成を行うモデルを提案する。具体的には、対話における状況を、システムの内部状態および発話権の状態を次元の一つとする多次元の心的状態として表現し、その変化の予測をプランニングの枠組みの中で行うことによってより状況に適した行為の決定を可能にする。第2章に述べたコミュニケーション調整機構の基本モデルとしての発話権管理モデル、さらに状態提示行為制御モデルを一般化したメタ発話生成モデルを、プラットフォーム上に実装し、評価実験を行った結果、対話の円滑さと効率の向上に効果があることが確認できた。

第6章には、本研究の総括を行い、残された問題点等を述べる。

本研究は音声対話インタフェースにおけるコミュニケーション調整機構の一般化を目指すものであり、基本となる発話権管理モデルから一般的なメタ発話生成モデルまでを構築し、実用を意図した汎用プラットフォームへの実装を行った。各モデルの形式化や汎用化の手法、さらにはプラットフォームのソフトウェア、収集した音声対話データおよびその分析ツールなどが本研究の成果としてあげられるが、従来工学的に扱われることの少なかったコミュニケーションの調整行為を音声対話インタフェースという実システムにおいて工学的に実現する取り組みそのものが本研究の価値であると考えられる。