

内 22-5

早稲田大学大学院理工学研究科

# 博士論文審査報告書

## 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

2002 年 6 月

音声による機械との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始められた。近年になってカーナビ、航空券予約サービスなどの実用システムも登場している。しかし現在なお数多くの研究課題が残されている。そのうちのひとつとして使用者の発話に対応して機械側の発話計画を行う、相互作用性向上のためのコミュニケーション調整の問題がある。

音声対話インタフェースは目的指向対話への利用が最も進んでいるが、目的指向対話でも目的達成のための直接的な情報のやりとりだけでなく、音声メディアの利点である相互作用性を高くするために、コミュニケーション調整の機能が重要となる。この機能を実現するアプローチとしては、音声に加えて顔の表情といった視覚情報や音楽等の聴覚情報を含んだ異種のモダリティを追加してコミュニケーション調整の役割を担わせる方法が提案されている。また、その有効性も報告されているが、電話のように音声情報のみをモダリティとするメディアにおいては適用できない。一方、こうしたコミュニケーション調整を学習によるモデル獲得の問題とするアプローチもあるが、原始的な相互作用を実現するにすぎず、問題解決の熟考的なプロセスと融合するのが困難となっているのが現状である。これに対して本研究は、問題解決を目的として構築した対話制御モデルにコミュニケーション調整機構を導入することにより、目的指向対話に起こり得る相互作用性の問題に対処している。

以下、各章の概要とその評価をまとめる。

第1章では、上述した研究の背景として、音声対話インタフェースの必要性と対話制御モデルの研究動向について述べている。また、その実現にあたって現状の技術が抱える問題を概観し、本研究の意義と位置付けを明らかにしている。

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している。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。本論文では、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で実現している。具体的には、人間同士の対話音声を収集して分析した結果に基づいて、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ている。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が提案されている。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確認されている。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ている。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点は、音声対話インタフェースにおける基本的な対話制御モデルを構築するうえで意義深い。

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が提案されている。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手法である。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されている。本章では、人間同士の対話についての評価指標特性を分析し、その結果に基づいて状態提示行為制御モデルの検討が行われている。対話制御モデルの構築においてその評価方法は重要な問題であり、その一つのアプローチを提案している点で価値が高い。

第3章で提案され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と位置付けられる。第4章と第5章では、調整行為全般についての制御を可能にする枠組みとして、まず音声対話インタフェース全体の汎用的な表現としてのプラットフォームが提案され、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的状態を導入することで、コミュニケーション調整機構の一般化がはかられている。第4章で述べられてい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している。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が実現可能になっている。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される。

第5章では、プランニングのレベルで対話における状況変化の予測に応じ、状況に適したメタ発話生成を行うモデルが提案されている。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している。構成としては、第2章に述べられ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、さらに第3章で述べられた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が、第4章で述べられたプラットフォーム上に実装されている。さらに評価実験により対話の円滑さと効率の向上に効果があることが確認されている。目的指向対話に限定されるが、発話計画の処理にコミュニケーション調整機構を導入することによって、問題解決のプロセスとコミュニケーション調整のプロセスをプランニングレベルのレベルで同一に扱えるようにしたこと、さらにコミュニケーション調整行為の範囲を拡大したことにおいて高く評価される。

第6章には、本論文の総括と残された問題点等が述べられている。

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したものであり、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が提案され、実用を意図した汎用プラットフォームへの実装が行われた。目的指向対話を扱う音声対話インタフェースにおいて重要な課題である使用者と機械間の相互作用性の向上に向けて、本研究での一連のコミュニケーション調整機構の提案は重要な寄与をなすものである。また、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も本研究の成果としてあげられるであろうが、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自体も価値あるものである。よって本論文は博士（情報科学）早稲田大学の論文として価値あるものと認める。

2002年5月

主査	早稲田大学	教授	工学博士（早稲田大学）	白井克彦
	早稲田大学	教授	工学博士（早稲田大学）	匂坂芳典
	早稲田大学	教授		寛捷彦
	早稲田大学	教授	工学博士（早稲田大学）	小林哲則