

# インタラクティブなプレゼンテーションでの ユーザ発話の自然な制限のための複数エージェントの利用

箕津真一郎<sup>†</sup> 中野幹生<sup>‡</sup> 船越孝太郎<sup>‡</sup> 竹内誉羽<sup>‡</sup>

長谷川雄二<sup>‡</sup> 土肥浩<sup>†</sup> 石塚満<sup>†</sup> 辻野広司<sup>‡</sup>

<sup>†</sup> 東京大学大学院情報理工学系研究科

<sup>‡</sup> (株) ホンダ・リサーチ・インスティテュート・ジャパン

## 1 はじめに

ロボットやアニメーションエージェントの応用のひとつとして、自動プレゼンテーションが注目されている。我々はこれまで、ユーザからの音声入力を受け付けてプレゼンテーションの内容を変える、インタラクティブなプレゼンテーションエージェント/ロボットシステムを開発してきた [1]。しかしながら、視聴者が、どのような発話内容で質問や要求を行ってよいか分からないという問題があった。視聴者の発話の認識・理解には、あらかじめ設計者が構築した認識・理解モデルを用いるため、視聴者の発話とそのモデルにマッチしなければ、システムはうまく理解できない。

この問題に対処するために、音声理解をロバストにして設計者の予測からはずれた発話でも理解できるようにすることや [2]、予測からはずれた発話を検出してヘルプを出す方法 [3] などが試みられている。

本稿では、これらとは異なるアプローチとして、プレゼンテーションを行うエージェントのほかに、もうひとつのエージェントを用い、質問や要求をデモンストレーションさせることにより、暗示的に、視聴者に発話の仕方を教える方法を提案する。この方法により、視聴者が質問・要求を行う際に、エージェントの発話と類似した発話を行う可能性が高くなり、結果として発話の認識・理解の精度が高まることが期待される。

## 2 インタラクティブなプレゼンテーションシステムにおける複数エージェントの利用

### 2.1 MPML-HR ver.3 システム

提案手法は、われわれが開発してきたインタラクティブプレゼンテーション記述言語 MPML-HR (Multimodal Presentation Markup Language for Humanoid Robots) ver.3 (以下 MPML-HR3) とその実行システム [1] を拡張することで実現している。MPML-HR3 実行システムは、MPML-HR3 で記述されたスクリプトに基づいて、ロボットやエージェントによるインタラ

クティブなプレゼンテーションを行うことができる。スライドを表示し、音声とジェスチャーを用いて、プレゼンテーションを進めて行く。視聴者は、インタラクション中に、「... を説明して」、「前に戻って」、「終わっていいよ」など、別の内容に遷移したりプレゼンテーションを終了したりするような命令を発話することができる。どのような発話に対して、どのような遷移を行うかは、あらかじめスクリプトの中に書かれている。発話パターンはネットワーク文法で記述する。質問・要求発話の認識結果の信頼度に応じて、そのまま遷移したり、問い返したり、無視したりすることができる。

### 2.2 複数エージェントの利用

視聴者に暗示的に発話の仕方を教えるために、エージェントをもう一体用い、質問・要求発話の実演を行わせる。これは、プレゼンテーションを行っている方のエージェントの発話にもう一体のエージェントが反応して質問・要求を行うような規則を記述することで可能にする。エージェントによる質問・要求が入力されると、視聴者からの質問・要求へ反応するのと同様に、プレゼンテーション内容の遷移を行う。このような拡張を行ったプレゼンテーション記述言語を MPML-HR ver.4 と呼ぶこととした。

## 3 予備実験

提案法の有効性の確認を目的に、予備的に被験者実験を行った。被験者がプレゼンテーションを見る際に、説明をするエージェントのみの場合と、それに加えて質問の手本を見せるエージェントを用いた場合に、被験者の質問・要求発話にどのような差異が生じるかを以下に示す方法で調べた。

方法 被験者は、日本語を母国語とする男性 12 名、女性 11 名 (年齢 20 歳-62 歳) である。実験では、2 体の Microsoft Agent を用いた。プレゼンテーションを行うエージェント (説明エージェント) と、質問の手本を見せるエージェント (質問エージェント) について、以下のような 2 条件を設定した。

質問エージェントなし 説明エージェントのみが登場してプレゼンテーションが行われる

質問エージェントあり 説明エージェントおよび質問エージェントが登場してプレゼンテーションが行われる

Using Multiple Agents to Naturally Restrict User Utterances in Interactive Presentations

Shinichiro MINOTSU<sup>†</sup>, Mikio NAKANO<sup>‡</sup>, Kotaro FUNAKOSHI<sup>‡</sup>, Johane TAKEUCHI<sup>‡</sup>, Yuji HASEGAWA<sup>‡</sup>, Hiroshi DOHI<sup>†</sup>, Mitsuru ISHIZUKA<sup>†</sup> and Hiroshi TSUJINO<sup>‡</sup>

<sup>†</sup> Graduate School of Information Science and Technology, The University of Tokyo

<sup>‡</sup> Honda Research Institute Japan Co., Ltd.

各被験者について、2回ずつプレゼンテーションを行った。被験者23名を[質問エージェントなし]、[質問エージェントあり]の順に行うグループと[質問エージェントあり]、[質問エージェントなし]の順に行うグループに、男女比及び人数がほぼ等しくなるよう配分した。今回は、各被験者について、最初のプレゼンテーションだけを分析対象とした。両条件とも提示する情報の総量(説明するページの数)が同じになるようにした。用いたプレゼンテーションの内容は、インターネットのウェブサイト(GooまたはGoogle)の説明である。質問エージェントありの条件では、質問エージェントが「～について教えてください」という文型で3回の質問を行った。この文型は、視聴者の発話の認識文法に含まれている。

手続き 被験者に、対話システムから情報を聞き出すという目的と、音声発話により対話システムが反応することを教示した。エージェントは音声発話を完全に理解することはできず、認識されない場合は話す内容やタイミングを変えるとよいということも教示した。また、質問エージェントの発話を聞いてから質問してもらうため、一通り説明が終わってから質問をするように教示した。

被験者に与えた課題は以下に示す通りである。

1. サイト名の由来を聞き取る
2. サービスの内容を2つ聞き出す

実験では、被験者の質問と考えられる発話が質問エージェントの発話と文型が似ているかどうか、及び音声認識に成功したかどうかを観察した。

仮説と予測 質問エージェントがある場合、質問エージェントの発話に影響され、質問エージェントの発話と同じ文型で話す被験者が多いことが予測される。被験者の最初の質問に注目することでエージェントの影響を最もよく観測することができると考えられる。また、被験者は自分の発話がうまく認識されたかどうかによって、発話パターンを変えようと考えられる。うまく認識された場合は、同じ発話パターンを繰り返し、それ以外の場合は、別のパターンを試すと考えられる。うまく認識できなかった場合、質問エージェントありの条件では、エージェントの質問発話パターンを試す可能性が高いと考えられる。

結果 被験者の最初の質問と、それがうまく認識されなかった場合のその後の質問を分析した結果を表1に示す。

ここで、「～教えてください」と被験者が発話した場合を「当該文型」とした。それ以外(「～教えて」も含む)を「別文型」とした。また被験者の発話に対し、その意図どおりにシステムが応答した場合を「認識」、その他を「誤認識」とする。

「当該文型・認識」は、最初の質問が当該文型であり、うまく認識された場合である。「別文型・認識」は、最初の質問が別文型であったがうまく認識された場合である。「別文型・誤認識→当該文型・認識」は、最初の質問が別文型だった場合、3つめまでに当該文型の質問をし、それがうまく認識できた場合である。「その他」の2人の被験者は、2人とも4つめまですべて別文型で質問し、4つめでうまく認識できた。

表 1: 被験者の質問の文型

文型のパターン	質問エージェントなし	質問エージェントあり
当該文型・認識	5	5
別文型・認識	3	4
別文型・誤認識 →当該文型・認識	1	3
その他	2	0
合計	11	12

考察 実験の結果、半数近くの被験者は質問エージェントの有無に関わらず、質問エージェントと同じ文型で最初の質問を行った。これは、質問エージェントの質問の文型が極めて一般的なものだったからと思われる。最初の質問が質問エージェントと異なる文型で、うまく認識できなかった場合(「別文型・誤認識→当該文型・認識」および「その他」)、質問エージェントありの条件では、3人とも3発話目までに質問エージェントの文型で質問した(「別文型→当該文型・認識」)のに対し、質問エージェントなしの条件では、はじめて認識する前に質問エージェント文型で質問したのは1人とどまった。これは、質問エージェントの発話が、「～教えてください」の文型に誘導した可能性がある。被験者の人数が少ないため、確定的なことは言えないが、提案手法の有効性を示唆する結果と言える。

#### 4 まとめ

本研究では、プレゼンテーションをエージェントが行う際に、質問の仕方の手本を見せるもう一体のエージェントを用いることで、視聴者の発話の自然な制限を行い、視聴者の発話の認識・理解を容易にする手法を提案し、実装した。予備実験の結果、質問エージェントの発話が被験者の質問の文型を、質問エージェントの発話の文型に誘導した可能性が示された。今後は、被験者数を増やしたり、認識精度をコントロールしたりすることにより効果的な実験を行い、提案手法の有効性を示して行きたいと考えている。

#### 参考文献

- [1] Y. Nishimura, et al.: "A Markup Language for Describing Interactive Humanoid Robot Presentations," Proc. IUI. 2007.
- [2] 政瀧他: 統計処理による入力文から中間表現への変換を用いた音声言語理解. 信学論 D-II, 82(2), 1999.
- [3] 福林他: 音声対話システムにおける発話パターンを教示するヘルプの動的生成. 人工知能学会 SLUD 研究会 A601-03, 2006.