

人間とエージェント間におけるフィラーの有用性の検証

岡崎 光太郎† 神田智子†

大阪工業大学情報科学部情報メディア学科†

1. はじめに

擬人化エージェントが一般ユーザに普及し始め、擬人化エージェントと一般ユーザとがインタラクションを行う機会が増えてきている[1]. 擬人化エージェントの中でも、人と対面で会話するエージェントを対話エージェントという。この対話エージェントと人とのインタラクションにおいても、人同士と同様に、バーバルコミュニケーション能力とノンバーバルコミュニケーション能力が必要とされる[1]. ノンバーバル行動をエージェントに実装することで、エージェントと人間間のインタラクションをより実際の人間間のインタラクションに近づけることができる。しかし、対話エージェント研究で開発されるエージェントは、完璧に流暢な発話を行うものが多い。実際の人間同士の対話では言いよどみや身体操作が含まれる。我々は、この言いよどみや身体操作を対話エージェントに実装することで、より人間らしいエージェントとのインタラクションが可能になるのではないかと考える。

この言いよどみは、フィラーと呼ばれ、本研究で着目したフィラーには、バーバルコミュニケーションにあたる「音声フィラー」、ノンバーバルコミュニケーションにあたる「行動フィラー」の2種がある[2]. 本研究では、この2つのフィラーが人とエージェントとのインタラクションにどう影響するのか検証する。

フィラーとは「それ自身命題内容を持たず、かつ他の発話と狭義の応答関係・接続関係を持たない、発話の一部を埋める音声現象」である[3]. 「えっと」や「あの一」などに代表されるように、それらはほぼ無意識に発せられ、「それを取り去っても伝達する文・談話の命題内容に変化を及ぼさないもの」とされている[4]. また会話中に生起する「よどみ」はこのような音声によるフィラーだけでなく、身振りにも存在すると考えられている[5]. 本研究では、このように発話の際に出現する動作的なよどみを「行動フィラー」と呼び、音声として発せられるフィラーをこれと区別して「音声フィラー」と呼ぶ。フィラーは人同士の対話に頻繁に用いられるものであり、『日本語話し言葉コーパス』では平均17語に1度の割合(総語数の約6%)で用いられており、高頻度の現象とされている[6].

フィラーに関連する研究では、人同士が対話しているビデオを「フィラーなし」、「音声フィラーあり」、「音声・行動フィラーあり」の3条件で印象評価を行った結果、「音声フィラーあり」は、「フィラーなし」よりも、対話の気まずさがやわらいだと感じること示された[2]. また同じく人同士が対話しているビデオを、沈黙あり、沈黙なしの2条件で印象評価を行った結果、沈黙ありの方は沈黙なしに比べ、より拒絶的な印象を抱かせた[7]. 上記2つの研究から本研究では、「無音区間にフィラーを実装していないエージェントよりも、無音区間にフィラーを実装したエージェントの方がよりエージェントに対する気まずさが軽減される」、「無音区間にフィラーを実装していないエージェントよりも、無音区間にフィラーを実装したエージェントの方がよりエージェントに対する拒絶的印象が軽減される。」という仮説をたて、人とエージェントとのインタラクション評価を行う。

2. 対話エージェント

エージェントのモデルおよびアニメーション作成に Poser7 を用い、それぞれの行動フィラー及び身体動作と発話中のアニメーションを作成した。発話音声は AITalk 声の職人, AITalk SDK を用いて作成、音声合成した。エージェントの対話システムの開発環境は Microsoft Visual Studio 2008, 開発言語は C++ である。対話システムはエージェントからの発話と、エージェントからユーザへの質問に対するユーザの回答待ち状態が遷移する。ユーザの回答により対話内容を分岐させ、エージェントの発話動画を再生する。

「フィラーなし」、「音声フィラーあり」、「音声・行動フィラーあり」の3条件で評価実験を行う。「音声フィラーあり」、「音声・行動フィラーあり」に実装した3種の音声フィラーは「えーっと」、「うーん」と、「そのー」である。音声フィラーは『日本語話し言葉コーパス』で「フィラー表現」とされている音声形態[6]から選択した。「音声・行動フィラーあり」に実装した3種の行動フィラーは、「横を向く」「上を向く」「口に手を当てる」であり、行動フィラーは音声フィラーと同時に開始されるよう実装した。「フィラーなし」には「音声フィラーあり」、「音声・行動フィラーあり」に実装した音声フィラー及び行動フィラーの部分に無音区間を作成した。

また行動フィラーとは別に、行動フィラーのみが目立たないようにするため、あいさつのときに「お辞儀する」、身体操作として「髪を触る」「顔を触る」といった動作をすべてのエージェントに実装した。行動フィラーを実装したエージェントのアニメーションを図1に示す。



図1. 行動フィラー「上を向く」「口に手を当てる」「横を向く」を実装したエージェント

3. 実験内容

3.1 手順

エージェントと実験参加者の対話方法は Wizard of Oz 法で行う。対話はエージェント主導で行われる。エージェントが質問し、実験参加者は、画面に表示された選択肢から発話によって回答し、研究担当者がシナリオに沿ってエージェントの対話内容を適宜操作することで対話を進めていく。対話内容はエージェントからの自己紹介や質問形式の日常的な対話を行う。対話時間は3分程度である。実験終了後にエージェントの印象評価に関するアンケートを行う。

3.2 印象評価

評価アンケートは、エージェントと対話に関する計19項目をリッカート尺度7段階で実験参加者に評価してもらった。評価アンケートのカテゴリの分類と質問項目を表1に示す。

また、気まずさの要素として「意識の統合性」「緊張感」「積極性」を[2], 拒絶的印象の要素として「受容性」「拒絶性」[7]を評価することで、気まずさと拒絶的印象を分析していく。

The Effects of Conversational Fillers by Virtual Agents

† Kotaro Okazaki, Tomoko Koda

† Department of Information Science, Osaka Institute of Technology

表1 アンケート項目

分類	質問内容
受容性	エージェントに理解してもらえていると感じた
	エージェントに信頼されていると感じた
拒絶性	エージェントによく思われていないと感じた
	エージェントにあまり相手にされていないように感じた
親しみ	エージェントと気軽に話すことができた
	エージェントとは親しみやすいと感じた
対話にたいするストレス	エージェントとの会話は不快であった
	エージェントとの会話は退屈だと感じた
対話のスムーズさ	エージェントとの会話は自然だと感じた
	エージェントと自分のお互いの意思が適切な速さで伝わった
対話にたいする好感度	エージェントとまた会話したいと感じた
	エージェントとの会話を楽しいと感じた
意識の統合性	エージェントは考えがまとまっていると感じた
	エージェントはゆったりしていると感じた
緊張感	エージェントが気楽にしているように感じた
	エージェントは意欲的であると感じた
積極性	エージェントが積極的であると感じた
	エージェントが積極的であると感じた

4. 分析結果

アンケート結果を3水準（フィラーなし、音声フィラーあり、音声・行動フィラーあり）の一元配置分散分析による分析、比較を行った結果を図2～図6に示す。分析結果からは以下の点が読み取れる。

- ① 拒絶的印象の要素である「受容性」、「拒絶性」において3条件間での有意差は見られなかった。(図2,3)
- ② 気まずさの要素である「意識の統合性」において3条件間での有意差は見られなかった。(図4)
- ③ 気まずさの要素である「緊張感」「積極性」音声フィラーありのエージェントよりもフィラーなし、音声・行動フィラーありのエージェントのほうが有意に高い結果となった。(図5, 6)

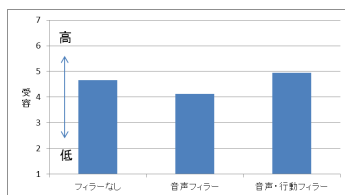


図2. 「受容性」の分析結果

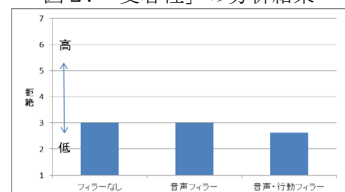


図3. 「拒絶性」の分析結果

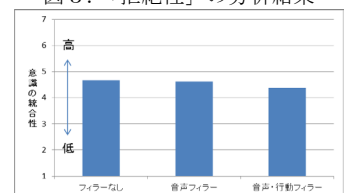


図4. 「意識の統合性」の分析結果

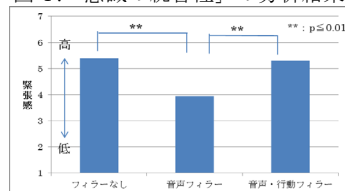


図5. 「緊張感」の分析結果

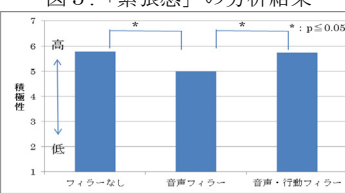


図6. 「積極性」の分析結果

5. 考察

分析結果①から、拒絶的印象の要素である「受容性」、「拒絶性」において3条件間での有意差は見られなかった。これは、エージェントから質問をする疑似対話方式のため、無音区間がぎこちない沈黙と捉えられたのではなく、エージェントからの発話待ちだとみなされたのではないかと考えられる。本実験における人とエージェント間のインタラクションにおいて、無音区間の存在による拒絶的印象を受けにくく、フィラーによる拒絶的印象の軽減は感じにくいのではないかと考えられる。

分析結果②から、気まずさの要素である「意識の統合性」において3条件の有意差は見られなかった。これは「意識の統合性」の質問内容である「エージェントは考えがまとまっていると感じた」において、3分という短い疑似対話では、エージェントの考えがまとまっているかまでは判断できず、人とエージェント間でのインタラクションにおける無音区間の存在による印象差が生まれず、フィラーが入ることによって考えのまとまっている印象は感じにくいのではないかと考えられる。

分析結果③から、気まずさの要素である「緊張感」「積極性」に関して、音声フィラーありのエージェントよりもフィラーなし、音声・行動フィラーありのエージェントのほうが有意に評価が高い結果となった。この結果から、人とエージェント間のインタラクションにおいて、音声フィラーと行動フィラーを同時に発生させることで、会話の中での言いよどんでいる状態を自然に表現できたが、音声フィラーのみではエージェントの言いよどみが不自然になり、「緊張感」にあたるエージェントに対するゆったりした印象や、「積極性」にあたるエージェントの対話に対する積極的な印象を感じにくかったのではないかと考えられる。またフィラーなしのエージェントと音声・行動フィラーありのエージェント間で「緊張感」「積極性」に関して有意差が出なかったのは、エージェントから質問をする疑似対話方式のため、無音区間がぎこちない沈黙と捉えられたのではなく、エージェントからの発話待ちだとみなされたのではないかと考えられ、フィラーの有無による「緊張感」にあたるエージェントに対するゆったりした印象や、「積極性」にあたるエージェントの対話に対する積極的な印象についての差を感じにくかったのではないかと考えられる。

6. おわりに

本研究では、フィラーなし、音声フィラーあり、音声・行動フィラーありの3条件のエージェントと実験参加者に対話してもらい、フィラーの有無による会話の気まずさの印象評価を行った。分析結果から仮説を支持する結果は得られなかったが、実験参加者はエージェントとの日常会話における無音区間に気まずさや拒絶的印象をあまり感じていなかったのではないかと考えられる。今後はよりシリアスな会話内容についてより長い対話を行う、より長い無音区間を発生させる、といった実験条件を変えることにより、人とエージェント間でも、無音区間における気まずさや拒絶的印象を感じ、無音区間におけるフィラーの効果を期待できるのではないかと考えられる。

参考文献

- [1] 山田誠二：人とロボットの<間>をデザインする，東京電機大学出版局(2007)
- [2] 佐々木寛紀，武川直樹，木村敦，徳永弘子：非円滑な発話交替時における沈黙の気まずさとフィラーの関係，信学技報，IEICE Technical Report HCS2011-13, HIP2011-13(2011-05)
- [3] 山根：日本語の談話におけるフィラー，くろしお出版(2002)
- [4] 中島：自然談話に現れるフィラー—自然談話録音資料に基づいて—，アジア・日本研究センター紀要，Vol.4, pp.1-23
- [5] 音声フィラーの抑制が動作フィラーの生起量に与える影響について，日本認知科学会 29 回大会(2012)
- [6] 前川喜久雄：『日本語話し言葉コーパス』の概要，日本語科学，15, 111-133(2004)
- [7] 宮崎弦太：会話場面で生じる沈黙の社会的拒絶としての機能，日本社会心理学会(2004)