

注意喚起の動作フィラーを用いた ファシリテーションエージェントの印象評価

Impression Evaluation of Facilitation Agents Using Attentional Motion Fillers

菅本祐也^{1*} 石井裕² 望月俊男³ 江木啓訓⁴ 横田真斗¹ 小畠弘暉²

Yuya Sugamoto¹ Yutaka Ishii² Toshio Mochizuki³ Hironori Egi⁴ Masato Yokota¹ Koki Kohata²

¹ 岡山県立大学大学院情報系工学研究科

¹ Graduate School of Computer Science and Systems Engineering,
Okayama Prefectural University

² 岡山県立大学情報工学部

² Faculty of Computer Science and Systems Engineering, Okayama Prefectural University

³ 専修大学ネットワーク情報学部

³ School of Network and Information, Senshu University

⁴ 電気通信大学大学院情報理工学研究科

⁴ Graduate School of Informatics and Engineering, The University of Electro-Communications

Abstract: In this research, we introduced an attentional motion fillers into a discussion support system using an interactive holographic agent, and conducted an evaluation experiment on the impression it gives to discussion participants. The attentional action fillers are intended to increase the inducement of the agent and to make it easier for the participants to accept the agent's instructions when the agent intervenes in a discussion. The setting of the experiment was a collaborative group discussion in which the agents played the roles of a facilitator and a timekeeper.

1 はじめに

グループでの協調的議論を行うために、ファシリテーターと呼称される、議論参加者の発言率や内容に応じて互いの議論を促進する振る舞いを行う進行役が置かれる場合がある。エージェントによるファシリテーションの先行研究では、人間によって行われるよりもロボットがファシリテーションすることで、参加者が中心となった積極的な議論を生み出す可能性があることが明らかにされている [1]。また、ロボットエージェントは高価で多数のシステム導入には限界があることから、プラスチック製の投影装置を用いてディスプレイ上のシステム動作映像を投影し、複数の参加者に対し3次元的なインタラクションが可能な疑似ホログラフィックエージェントの開発も行われている [2]。WoZ (Wizard of Oz) 法によるシステムを用いた評価実験によりエージェントの有効性を検証し、議論参加者の音声集約システム [3] と結合してファシリテーションシステムを

構築している。対面場面における学習者中心の学習環境デザインを志向する場合、エージェントと学習者の関係性が、エージェントに対する認識の在り方に影響があることが示されている [1]。発話開始の聞き取りやすさに着目し音声フィラー機能を追加したシステム [2] に音声フィラー機能を追加したものをを用いた研究では、音声フィラーあり、なしにかかわらず、議論が進むにつれてエージェントへの興味が低下し、指示を受け入れる回数が減少する傾向が確認された。会話におけるフィラーとは「えっと」や「あのー」などの言いよどみで「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・修飾関係にない、発話の一部を埋める音声現象」 [4]、「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」 [5] などと定義されている。会話中において生起する「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おり [6]、動作フィラーを用いた研究も行われている。さらに、人が対面して行うコミュニケーションでは、言葉の意味や内容によるバーバル情報だけではなく、身振り・手振り、うなず

*連絡先：岡山県立大学
〒719-1197 岡山県総社市窪木 111
E-mail: sk622031@cse.oka-pu.ac.jp

きや表情などの視覚情報、声の音調や会話速度によってもたらされるノンバーバル情報が重要な役割を果たしており [7], 協調的議論においてもノンバーバル情報の動作フィラーが有効である可能性が示唆されている。

本研究では、エージェント介入時における発話内容の受け入れやすさと誘目性を向上させることを目的とし、話を振る対象に目線を合わせ姿勢を前面に出す動作を、介入時の動作フィラーとして表現する機能を追加した。評価実験を行い、フィラーの有効性を検証した。

2 対話型ホログラフィックエージェント

本研究で用いたシステムは、先行研究の対話型ホログラフィックエージェントシステム [2] をもとに、動作フィラー機能とタイムキーパー機能の追加を行った。先行研究のシステムは議論における共同学習を支援することを目的として開発されている。またエージェントは身体動作（身体全体の回転、腕部、頭部動作）と音声によるファシリテーションを行うことで、議論参加者の発話を推進し、議論の活性化を促進している。

2.1 動作フィラー機能

エージェントは発話率などの条件で介入するため、話を遮るかたちで介入する場合がある。このとき、議論参加者は指示を聞く態勢をとるのに時間がかかるため、提示した内容がうまく聞き取れないことがある。また発話開始の聞き取りやすさを向上させるために、先行研究のシステムに音声フィラーを用いた予備実験を行った結果、エージェントを元々見ていない場合は、介入が行われても参加者が指示を聞き入れない傾向が確認された。

そこで発話内容の受け入れやすさと誘目性を向上させることを目的とし、注目を促す動作によるフィラー機能を追加した。動作フィラーの挙動は、議論において発言するとき体が出ることを参考に、図1の流れで図2に示すようにエージェントの大きさを拡大することで実現した。また、動作に対して発声が0.3秒程度遅延するあいさつが好ましいことが明らかにされており [8], より自然な動作を行うために、エージェントが発話するときの動作に対して発話を0.3秒遅延させた。

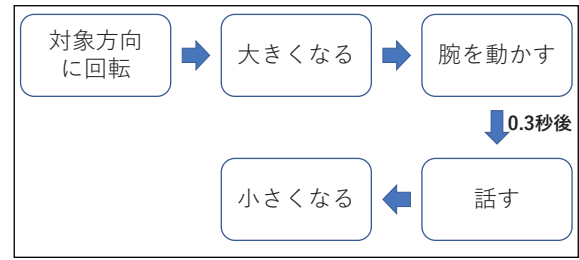


図1: 動作フィラーの流れ



(a) フィラー前

(b) フィラー後

図2: 動作フィラー時の変化

2.2 音声によるタイムキーパー機能

エージェントの指示に意識を向けさせ、参加者が指示を受け入れやすい状態を生み出すために、音声によるタイムキーパー機能を追加した。これにより図3のコンセプト議論の経過時間に合わせた介入と経過時間の告知を合成音声を用いて提示する。音声での指示内容については、表1に示すように議論の時間と意見の掘り下げ、集約といったファシリテータとしての指示を設定した。そのためタイムキーパー機能の介入は、音声集約システム [3] からファシリテータとして介入する指令があるときに、代わりにタイムキーパー機能の介入を行う。本研究では議論時間を10分で設定しているため、タイムキーパー機能の介入は0分、2分、5分、8分経過後に行った。

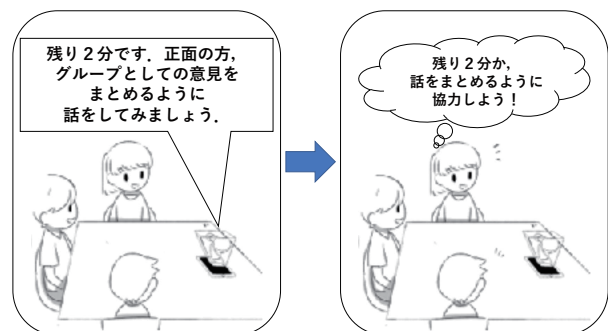


図3: タイムキーパー機能のコンセプト

表 1: タイムキーパー機能のプロンプト一覧

経過時間	指示の内容
0分	今回の議論ではファシリテーターとタイムキーパーの役割を担います。
2分	現在2分経過しました。私の正面の方、今までの意見を何か掘り下げてみませんか？
5分	現在半分が経過しました。 皆さん書いてあるメモを見て、何か気になることがあればそれについて話をしてみましょう。
8分	残り2分です。正面の方、グループとしての結論をまとめるように話をしてみましょう。

3 評価実験

3.1 実験概要

2章のシステムを用いて、グループでの協調的議論におけるエージェントの指示の受け入れやすさと誘目性に対する効果の比較を行うために、エージェント介入時の動作フィルターあり、なしの2パターンで印象評価実験を行った。議論前に実験参加者に対して、エージェントの役割と指示の例について説明を行った。実験参加者は3人1組で、エージェントを交えて「幸せとは何か」「大人とは何か」について各10分間議論した。また、エージェント正面に着席している実験参加者を書記とし、議論の内容についてメモをとることを指示した。各議論終了後に23項目の7段階評価アンケートを行い、各条件は組ごとに順序と座席を入れ替えて提示した。評価する項目については3.3節で述べる。以上の条件で実験を行い、評価結果をもとにフィルターによる効果と役割に対する印象を分析した。実験参加者は、10～20代学生18名で、うち男性14名、女性4名であった。

3.2 実験環境

図4に協調的議論の実験を行った状況の全体図を示す。プラスチック製の投影装置を取り付けたモニターを1台設置し、そこに投影されるエージェントを中央として3方向に各実験参加者を配置した。投影装置の後ろにはスピーカーが配置され、エージェントの音声流れる。また、エージェントが映し出される投影装置内にWebカメラを3台取り付けており、各実験参加者がエージェントを見ているかを録画した。議論全体の様子は、三脚付きのビデオカメラを中央の実験参加者の後ろ側に配置し、録画した。

実験参加者には Raspberry Pi と指向性マイク付きヘッドセットを各1台ずつ用意し、書記を割り当てた人には追加でメモ用の画用紙とマジックペンを用意した。実験参加者はヘッドセットを首にかけ、口元にマイクを持っていくことで音声集約システム [3] で使用する音声の取得を行った。実験者側にはノートパソコンをカメラの録画用とシステムの動作用に2台配置した。また、実験参加者と同室で実験の説明・指示と実験状況の観察を行った。

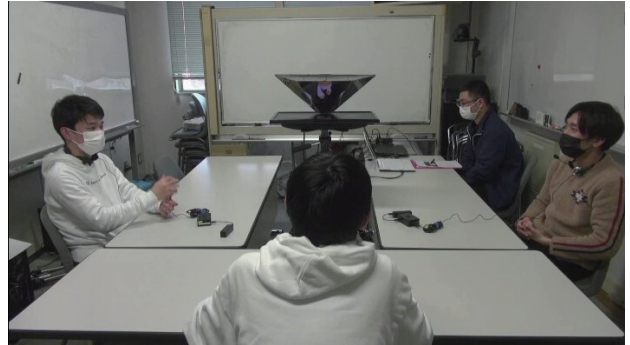


図 4: 協調的議論の実験風景

3.3 評価項目

表2に、実験で評価した項目を示す。Q1ではエージェントの受容性についての問いを6項目、Q2ではエージェントの介入についての問いを5項目、Q3では議論全体についての問いを4項目、Q4では議論に対する自分の振る舞いについての問いを3項目、Q5では実験環境についての問いを2項目、Q6では追加した動作フィルター機能とタイムキーパー機能の評価についての問いを3項目用意した。これらを用いて7段階の官能評価を行った。またグループとしての結論を記述する項目と自由記述を用意し、結論とその他の印象や感想についても記入する形式とした。

表 2: 評価項目

設問番号	設問内容
Q1-1	エージェントの音声の流れたときに驚いた
Q1-2	エージェントの発言内容の聞き取りは容易であった
Q1-3	エージェントのキャラクターが見づらかった
Q1-4	話し合い中にエージェントの存在を意識した
Q1-5	エージェントは自分のほうを向いているように感じた
Q1-6	エージェントも話し合いに参加しているように感じた
Q2-1	話をするように促される以前から、自分から発言しようと考えていた
Q2-2	話をするように促されたことがきっかけで発言を行った
Q2-3	話をするように促されたことで不快に感じた
Q2-4	話をするように促された音声の内容は自然であった
Q2-5	話をするように促された音声のタイミングは自然であった
Q3-1	議論のテーマについてじっくり話されていた
Q3-2	議論は活発だった
Q3-3	幅広い観点で話し合うことができた
Q3-4	議論全体を通して満足できる内容だった
Q4-1	議論のテーマについて自分はじっくり話せた
Q4-2	議論に率先して参加した
Q4-3	幅広い観点で自分は意見を出すことができた
Q5-1	マイクを装着することにより、議論に支障があった
Q5-2	議題の影響により議論の難易度が変化した
Q6-1	会議の残り時間を意識して議論を行った
Q6-2	エージェントの動きが気になった
Q6-3	エージェントの行動は不快であった

3.4 結果

23項目での7段階官能評価の結果について Wilcoxon の符号順位検定を行ったものを、それぞれ図5~10に示す。評価実験の結果、Q1-3「エージェントのキャラクターが見づらかった」の項目が有意水準5%で有意差が認められた。

また Web カメラで録画した映像から、エージェントの介入回数と介入時に実験参加者がエージェントを視認した回数について計数し、比較を行った結果を表3に示す。フィルターありの平均がフィルターなしに比べて11.6%高く、介入時にエージェントを見ている傾向が示された。またエージェントの介入時に指示を受け入れた回数の平均は、タイムキーパーとしての指示が他の指示に比べて22.5%高く、タイムキーパーの指示を受け入れている傾向が示された。

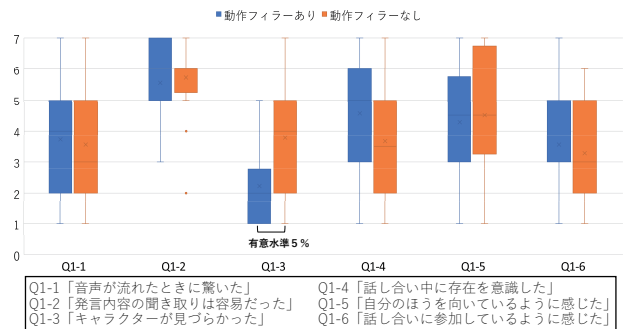


図 5: Q1「受容性」の官能評価結果

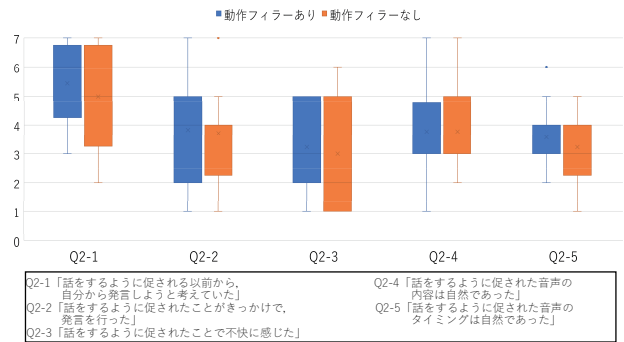


図 6: Q2「介入」の官能評価結果

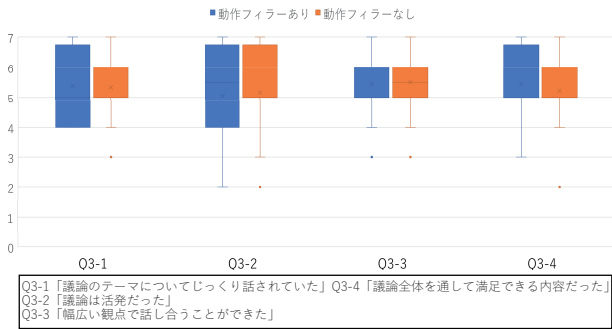


図 7: Q3 「議論全体」の官能評価結果

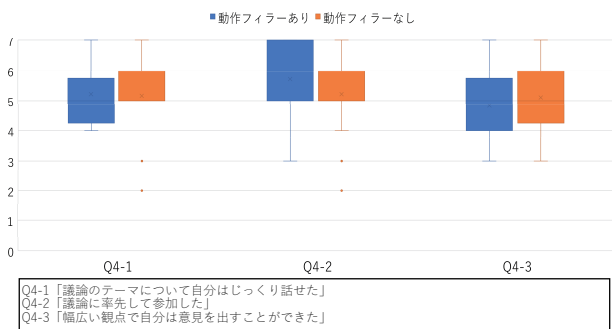


図 8: Q4 「自分のふるまい」の官能評価結果

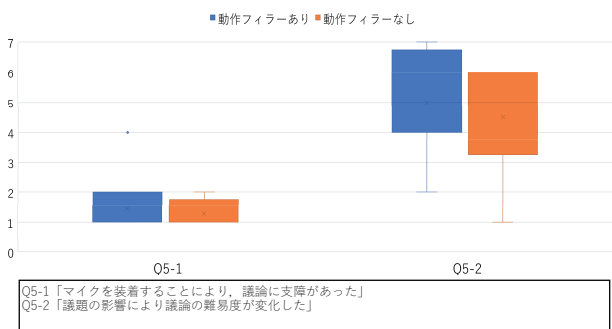


図 9: Q5 「実験環境」の官能評価結果

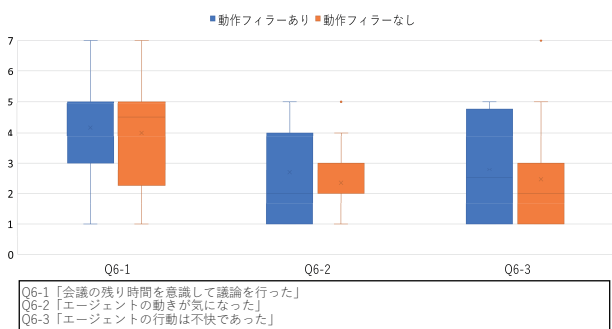


図 10: Q6 「追加システムの評価」の官能評価結果

表 3: 視認回数の比較

	動作フィルターあり	動作フィルターなし
視認回数	90.4 %	78.8 %
受け入れ回数	82.8 %	60.3 %

3.5 考察

動作フィルターあり、なしで評価した結果について、Q1「受容性」については、Q1-3よりエージェントのキャラクターの見やすさについて、有意水準5%で有意差があった。これは動作フィルターによりエージェントの大きさに差があることで、キャラクター自体が見やすくなったと考えられる。しかしQ1-1, Q1-2より、動作フィルターによる評価の差は見られなかった。これは指示が入る前にエージェントを見ていることが少ないため、評価に差がみられなかったと考える。ただWebカメラの映像中に、実験参加者がエージェントを見ているときに動作フィルターが行われた際、驚くような仕草を取りその後指示を受け入れた例がいくつか見られた。

Q2「介入」については、エージェントの介入タイミングについて評価が低い傾向であった。これは原因として、音声取得からエージェントの指示が行われるまでにタイムラグがあり、介入のタイミングがずれてしまったこと、エージェントの発話時に実験参加者は発話をしないことが多く、発話率の低いときの指示が出ることがあげられる。そのため、自由記述においても「話を遮られることが多々あった」といった記述が複数見られ、音声のタイミングについては改善の余地があることが確認された。

Q3「議論全体」、Q4「自分のふるまい」については条件間に差はなかったものの、全体的には満足のいく内容であったと評価されている。

Q5「実験環境」について条件間に差はなく、マイクの装着は特に支障がなかったと評価された。また、議題の影響で難易度が変化したと感じる者が多くみられた。

Q6「追加システムの評価」について、自由記述で「途中の何分経過ですは非常にありがたかった」といった意見が見られたが、会議の残り時間を気にしたかどうかには評価にばらつきがあった。これは、エージェントが時間を管理することで時間を気にする必要がないと感じる者と、エージェントが時間を管理することで時間を気にして議論を進めようとする者で印象が変わることが原因だと考えられる。またエージェントの行動に関しては、特に気にならず不快ではなかったと評価された。

Webカメラで録画した映像の分析については、動作フィルターありが動作フィルターなしに比べて見ている頻度が11.6%高かった。このことからアンケートでの印

象評価ではあまり差は出なかったが、エージェントの大きさの違いにより、実験参加者の視界に入りやすくなることで見る頻度に差が出たと考える。また、タイムキーパーとしての指示が他に比べて受け入れられている傾向が確認された。

4 おわりに

本研究では、対話型ホログラフィックエージェントシステムに注目を促す動作によるフィルター機能と音声によるタイムキーパー機能を追加し、音声集約システムを用いて印象評価実験を行った。その結果エージェントの見やすさについては、動作フィルターとして大きくなることの有効性が示され、エージェントと学習者の関係性構築において、エージェントの受容性が向上する傾向が見られた。しかし、エージェント介入時の音声指示が入る前からエージェントを見ていることが少なかったため、動作フィルター自体の有効性を示すにはさらなる工夫が必要である。またタイムキーパーの役割を追加で与えたことで、エージェントの指示を聞き入れやすくなった。このことから、時間経過などの特定の役割を持った内容の指示を行うことでエージェントの指示を聞き入れやすくなり、議論の発展に繋がると考える。

今後は、本研究で得られた結果をもとにシステムの改善を行う。また、エージェントの受容性を高めるための異なるアプローチや、議論での役割について比較を行う。これにより、協調的議論における対話型ホログラフィックエージェントシステムのさらなる開発を進めていく。

謝辞

本研究は、科学研究費補助金 (21K18527) の支援を受けている。

参考文献

- [1] Toshio Mochizuki, Yoshitaka Mitate, Yoshikazu Tateno, Takehiro Wakimoto, Yuko Miyata, Jun Nakahara, and Nahomi Miyake. Robot as a learning partner for promoting proactive discussion in peer groups: a case study for career development. In *International Conference on Computers in Education*, L.-H. Wong, Ed. Asia-Pacific Society for Computers in Education, 2013.
- [2] Ayane Hisatomi, Yutaka Ishii, Toshio Mochizuki, Hironori Egi, Yoshihiko Kubota, and Hiroshi Kato. Development of a prototype of face-to-face conversational holographic agent for encouraging co-regulation of learning. In *Proceedings of the 7th International Conference on Human-Agent Interaction*, pp. 308–310, 2019.
- [3] Naruaki Ishikawa, Taishi Okazawa, and Hironori Egi. Diana-ad: Dialog analysis for adjusting duration during face-to-face collaborative discussion. In *International Conference on Collaboration and Technology*, pp. 212–221. Springer, 2019.
- [4] 山根智恵. 日本語の談話におけるフィルター. くろしお出版, 2002.
- [5] 中嶋悦子. 自然談話に現れるフィルター – 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, pp. 1–23, 2009.
- [6] 阪田真己子, 鈴木紀子, 桑井春菜. 音声フィルターの抑制が動作フィルターの生起量に与える影響について. 2012年度日本認知科学会第29回大会.
- [7] 黒川隆夫. ヒューマンコミュニケーション工学シリーズ-ノンバーバルインタフェース-. オーム社, 1994.
- [8] 山本倫也, 渡辺富夫. ロボットとのあいさつインタラクションにおける動作に対する発声遅延の効果. ヒューマンインタフェース学会論文誌, Vol. 16, No. 3, pp. 343–350, 2004.