

エージェントとの会話におけるトップダウン処理とボトムアップ処理: 眼球運動と印象評価に着目した実験的検討

Top-down processing and bottom-up processing in conversation with the agent: Experimental study focusing on eye movement and impression

星田 雅弘[†], 田村 昌彦^{††}, 林 勇吾^{†††}

Masahiro Hoshida, Masahiko Tamura, Yugo Hayashi

[†] 立命館大学文学研究科, ^{††} 立命館大学文学部, ^{†††} 立命館大学総合心理学部

Ritsumeikan University Graduate School of Letters[†], Ritsumeikan University College of Literature^{††},

Ritsumeikan University College of Comprehensive Psychology^{†††}

lt0610fs@ed.ritsumei.ac.jp[†]

Abstract

In this study, we investigated the effects of two types of cognitive processing, the top-down processing and the bottom-up processing during communication with an agent. Especially, we investigated participant's impression to the agent and their eye movement. It was hypothesized that these two processing works on the basis of the agent's behavior respectively. We manipulated the agent's intelligence by the experimenter's instruction as top-down processing, and the degree of agent's intellectual behavior, as bottom-up processing. The results show that people feel agent's humanity depending on the each of the two processing, but feel agent's intelligence depending on the combination of the two processing. Results on eye movement show that people pay attention to an agent depending on its behavior, and expectation about an agent is not effective.

Keywords — Agent, Top-down, Bottom-up, Attentive action, Impression evaluation

1. はじめに

我々は、日常的に様々な人と接触し、コミュニケーションを行っている。その際、相手がどのような人物なのかを推論する方略として、トップダウン処理とボトムアップ処理の2種類の方略がある。トップダウン処理とは、相手に関する既有知識に基づいて行う情報処理である。ここでの既有知識とは、行動と特性の関係に関する知識や、社会的カテゴリーや役割によって区分される集団と成員に関する知識（ステレオタイプ等）等がある。例えば、同じ攻撃的な行為でも、行為

者が白人である場合よりも黒人である場合の方が、より攻撃的な印象を与えることが示されている [1]。本論文では、このように既有知識から相手のことを推論する処理のことをトップダウン処理と呼ぶ。

一方、ボトムアップ処理とは、相手から表出される情報に基づいて行う情報処理である。例えば、ある人が自分に対して親切に接してくれたとする。その結果、その人物に対して親切な印象が形成されるだろう。その後、またその人が自分に対して親切な振舞いをしてくれたとする。すると、親切な印象は強められるだろう。しかし、その人物が自分に対して攻撃的な態度を示した場合、一旦形成された親切な印象が改められることがあるだろう。本論文では、このように振舞いから相手のことを推論する処理のことをボトムアップ処理と呼ぶ。

1.1 会話における認知情報処理

我々は人とのインタラクションだけでなく、エージェントとのインタラクションにおいても、前節で取り上げた2種類の認知情報処理が影響することが明らかになっている。例えば、林・三輪 [2] は、(1) 既有知識の要因と (2) 相手の種類の要因の2要因がコミュニケーション特性に与える影響を実験的に検討した。(1) 既有知識としては、相手が人間であるかコンピュータであるかという2種類の条件を設定して操作した。一方で、(2) 相手の種類としては、実際の相手が人間かコンピュータかどうかという2種類の条件を設定して操作した。それぞれの条件下で実験参加者は、ネットワーク上に接続された他者とテキストチャットを用い

て規則発見課題を行った。その際、コミュニケーション時の印象評価にどのような影響を及ぼすのかを検討した。その結果、(1) 既有知識の要因はコミュニケーション時における対人緊張と親和感情といった感情特性に影響を与え、(2) 相手の種類の要因はコミュニケーション時における感情特性に加え、情報伝達といった認知特性にも影響を与えることが明らかになった。これは、本論文で示すところのトップダウン処理とボトムアップ処理がテキストベースの他者との会話において現れていたと考えることができる。ここで着目すべき点は、実際の相手の種類によって情報伝達のコミュニケーション特性が変化したことが示すように、相手の振舞いがコミュニケーション特性に大きく影響している点である。これは、エージェントとの会話において、ボトムアップ処理が重要であることを示しているといえる。

一方で、人間同士の会話における言語表現の利用に関しては、相手との親密度や話題の内容に対する知識の共有度の大きさによって、利用する言語表現が変わることが明らかになっている。例えば、Wuら [3] は、既有知識の要因が言語表現の利用にどのような影響を及ぼすのかを実験的に検討した。既有知識としては、話者がお互いに知っている単語が多いか少ないかという2種類の条件を設定して操作している。実験参加者は、それぞれの条件下で無意味図形を伝える課題を行った。その際、相手との知識の共有度が言語表現の利用にどのような影響を及ぼすのかを検討した。その結果、お互い知っている単語が少ないと考えている場合よりも、多いと考えている場合の方が、特定の単語で伝える傾向がみられた。このことは、相手との知識の共有度が高いか低いかという既有知識が、言語表現の利用に影響することを示している。すなわち、人間同士の会話における言語表現の利用・選択の段階において、トップダウン処理が影響していることが示唆される。

エージェントとの会話においても、既有知識が言語表現の利用に影響を与えることが明らかになっている。例えば、エージェントと一緒に同じ対象を単語で伝え合う場面を設定し、エージェントの性能が高いか低いかという認識の要因を検討した研究では、性能が高いと認識している場合よりも、低いと認識している場合のほうが、相手（エージェント）と同じ単語を用いる傾向がみられた [4]。これは、我々はインタラクションの相手であるエージェントが高性能であるという既有知識から、知的な印象を形成し、多様な言語が通じると推論していたと考えられる。このことは、エージェ

ントとの会話における言語表現の利用・選択の段階においても、トップダウン処理が影響していることを示唆するものである。

しかし、林・三輪 [2] の先行研究では、規則発見課題を用いた検討を行っており、言語の利用に関しては十分な検討が行われていない。更に、用いられた課題では、テキストチャットによってコミュニケーションが行われていたため、コミュニケーションにおいて重要な非言語情報が無視されていたことが考えられる。また、Pearsonら先行研究では、エージェントに対する言語表現の利用に関して、トップダウン処理の要因だけが検討されており、ボトムアップ処理の要因に関しては検討されていない。そこで、本研究では、言語選択課題を用いて、エージェントとの会話においてトップダウン処理とボトムアップ処理がどのように影響するのかを詳細に検討する。

1.2 エージェントに対する印象評価

対人印象評定では、視線の動きが重要となることが知られている。例えば、相手を見つめる時間の割合（凝視量）がその人物の力量や好悪の印象に影響するとする Cookら [5] および Argyleら [6] の研究や、好悪の印象に凝視量と発言の親和性の交互作用が見られるとする Ellsworthらの研究 [7]、キョロキョロしたり目を伏せたりすると不安そうに見えるとする Waxerの研究 [8] などがある。このように、視線はそれを表出した人物に対する印象形成に深く関わっている。また、これまでの研究から、対人認知場面と同様に、エージェントに対しても視線が印象形成に影響を与えることが指摘されている。神田ら [9] は、ロボットの視線制御を操作し、人間が受ける印象を実験的に検討した。印象評価に関しては、4因子（親近性因子/愉快性因子/活動性因子/性能評価性因子）を用いて測定した。その結果、親近性因子、愉快性因子、活動性因子においては視線制御の影響がみられず、性能評価性因子においては、視線制御を伴う場合のほうが、視線制御を伴わない場合よりも得点が低かった。すなわち、視線制御をすとかえって知的に感じられなかった。これは、発話速度が遅いなど、ロボットの性能がそもそもあまり高くなかったため、視線制御を行った結果、余計に性能が低いと判断されてしまったのではないかと考えられた。人間はシステムから表出される情報（外見や振舞い）からメンタルモデルを構築し、行動を予測する [10] ことから、ロボットが視線制御を伴うことによって知的な印象を形成し、高度な動きを予測して

いたと考えられる。ところが、高い期待を抱いている場合に実際の行動が期待外れであったため、知的な印象は余計に下がってしまったと考えられる。このように、期待外れである行動によって印象が悪化する現象を適応ギャップと呼んでいる [11]。知的な印象を形成するためには、視線以外の要因も適切に操作することにより、知的な印象を一貫させる必要があると考えられる。例えば、神田ら [9] の実験ではロボットの発話速度が遅かったことが問題であった。実際、人間に対する印象形成では、発話速度が遅い場合よりも、速い場合のほうが知的に感じられることが明らかになっている [12]。よって、発話速度も知的な印象の一因として操作することが解決策となる。本研究では、知性を伴う振舞いとして、視線制御以外の要因がエージェントの印象にどのような影響を与えるのかを検討する。

1.3 エージェントへの注視行動

我々は会話場面において、話し手は、聞き手が関心をもって会話に参加していることを聞き手の視線から察知し、関心がないと判断できれば話題の内容を変える等、会話の内容を調整している。一般的には、聞き手から話し手への注視行動は、話し手から聞き手への注視行動よりも多いといわれている [13]。聞き手が話し手に注意を向けることは、聞き手による会話への参加意思を示し、これが話し手へのフィードバックとなる。従って、対面の会話における視線行動はコミュニケーションを促進させるために重要な役割を担っている。

上記は対人場面における注視行動の役割であるが、エージェントに対する場面でも注視行動は一定の役割をもつ。例えば、エージェントに視線制御を実装することにより、人間同士に近い会話を実現できることが明らかになっている [14]。石井・中野 [15] は、エージェントの視線制御が会話にどのような影響を与えるのかを検討した。実験参加者が逸脱した会話参加態度であると推定されたときに、その実験参加者の様子をうかがう発話をする場合と、定期的の実験参加者の様子をうかがう発話をする場合の2条件を設定し、インタラクションを行った。その結果、ユーザと態度に応じて視線制御を行うほうが、一定の視線制御を行う場合よりも、ユーザの逸脱した態度が減少した。さらに、エージェントへの好感度が増し、人間らしく知的に感じられていた上に会話も円滑に感じられていた。このことから、エージェントとの会話においても、視線制御は円滑なコミュニケーションを促進させるといえる。

一般に、エージェントに対する印象やコミュニケーションのパフォーマンスは課題後にアンケートによって測定されることが多い。それに対し、注視行動はリアルタイムで測定できるため、会話のパフォーマンスや相手（エージェント）への好感度を正確に測定することができると考えられる。そこで、本研究では、人間の注視行動に応じたフィードバックをエージェントの振舞いとして操作し、人間の注視行動を測定する。

1.4 本研究の目的と仮説

以上を踏まえ、本研究では以下の2点を検討する。

1. トップダウン処理とボトムアップ処理がエージェントの印象評価にどのように影響するのかを明らかにする。

2. トップダウン処理とボトムアップ処理がエージェントへの注視行動にどのように影響するのかを明らかにする。

印象評価に関しては、対人認知処理にはトップダウン処理とボトムアップ処理の2種類の方略があることから、以下の仮説が導かれる。

H-1-a: エージェントに対して知的であるという既有知識があると、人間らしく知的に感じられ、実際の振舞いは影響しない

H-1-b: エージェントが知的に振舞うことによって、人間らしく知的に感じられ、既有知識は影響しない

H-1-c: エージェントの振舞いが顕著になると、既有知識がある場合でも、その影響は弱まり、知的な振舞いの増大によって人間らしく知的に感じられる

注視行動に関しても、エージェントに対して好感を抱き、会話が円滑に進んでいると注視行動が増えることから、以下の仮説が導かれる。

H-2-a: エージェントに対して知的であるという既有知識があると、エージェントへの注視行動は増え、実際の振舞いは影響しない

H-2-b: エージェントが知的に振舞うほど、エージェントへの注視行動は増え、既有知識は影響しない

H-2-c: エージェントの振舞いが顕著になると、既有知識がある場合でも、その影響は弱まり、エージェントへの注視行動は増える

2. 方法

2.1 実験参加者

大学生 90 名（男性 58 名、女性 32 名、平均年齢 21.12 歳）が参加した。

2.2 実験計画

実験は、2 要因の参加者間要因の計画により実施された。ここでは、トップダウン処理の要因とボトムアップ処理の要因を実験的に操作した。トップダウン処理は、エージェントへの先行知識の要因を操作した。具体的には、教示によって、エージェントが知的であると思わせるかどうかを操作した。以後、このトップダウン処理に関する操作を教示の操作と呼ぶ。ボトムアップ処理は、エージェントの振舞いを操作した。ボトムアップ処理に関しては、エージェントの振舞いの変化の大きさを3水準(統制条件/機能低条件/機能高条件)に設定した(表1参照)。これらの振舞いが変化することでボトムアップ処理が優勢になることが予想され、特に統制条件、機能低条件、機能高条件という順で、より知的かつ人間らしく感じられ、エージェントに対して注意を向けるようになることが期待される。以後このボトムアップ処理に関する操作を振舞いの操作と呼ぶ。

表1 振舞い要因の各条件の設定

	統制条件	機能低条件	機能高条件
(1) 視線	動作なし	追従なし	追従・瞬き
(2) 感情	なし	読み取りなし	読み取りあり
(3) 伝達	文字	音声(遅い)	音声(速い)
(4) 応答	必ず正解	必ず正解	正誤判定

振舞い要因の各条件の設定について詳細に述べる。振舞いは、(1) 視線、(2) 感情、(3) 伝達、(4) 応答の4種類に焦点を当てて設定した。まず、(1) 視線では、会話において相手を見ること(相手注視)や、共有する対象を見ること(共同注視)はロバストな会話につながることを示されている[16]。また、瞬きをある程度したほうが、全くしないより知的に感じられる[17]。(2) 感情では、我々は、類似した態度をとる他者に対して好意を抱く傾向があり、親和動機として知られている[18]。親和動機とは、自分に似ていたり、好意を持ってくれたりする人に対して、協力し、好意に報いることを求める欲求である。エージェントに対してもこのような親和的反応を引き起こすことが明らかになっている[19]。また、(3) 伝達では、発話速度が速いほうが知的に感じられる[12]。最後に、(4) 応答では、対応づけと自然なフィードバックが重要である[10]。

これらを実験条件ごとに異なる操作を設定した。以下に実験条件ごとに述べる。なお、実験では、統制条件、機能低条件、機能高条件の順に実施した。

2.2.1 統制条件

(1)(2)に関しては、エージェントは静止画であり、視線制御も感情表出も伴わなかった。(3)に関しては、伝達は文字によるものだった。(4)に関しては、応答は実験参加者の入力言語にかかわらず同じフィードバックを返した。

2.2.2 機能低条件

エージェントはアニメーションで動作した。(1)に関しては、実験参加者に話しかけるときは正面を向き(相手注視)、画像に話すときはその画像を見る(共同注視)といったように、会話に沿った視線の動きを行った。(2)に関しては、動作と音声によって一定の喜怒哀楽の感情表出を行った。(3)に関しては、修正 BSD ライセンスでオープンソースとして公開されているプログラム[20]を用いて作成し、発話速度は8.0(morae/sec)に設定した。(4)に関しては、入力言語にかかわらず音声にて同じフィードバックを返した。

2.2.3 機能高条件

機能低条件の機能に加え、PCに設置されているWebカメラを通して実験参加者の顔の位置や感情を読み取り、読み取り結果に基づいてエージェントが音声情報でフィードバックを返した。顔の位置や感情の読み取りには、clmtracker[21]を実験プログラムに組み込んだ。この機能により、(1)に関しては、実験参加者の顔の一部が一定の範囲外に検出されると、画面を見るように音声で催促した。その範囲は、全体400pix × 320pxのうち、座標で左上が(20, 20)、右上が(380, 20)、左下が(20, 300)、右下が(380, 300)とした。なお、ディスプレイ上でのエージェントの座標は左上が(430, 50)、右上が(930, 50)、左下が(430, 550)、右下が(930, 550)であった。また、瞬きを行うように設定した。(2)に関しては、表情から感情を予測する指標[22]に基づき、4種類の感情(angry/sad/surprised/happy)の強さを算出し、この値に基づいてフィードバックを返した。具体的には、実験参加者の一定間隔での各感情値の平均値を求め、最も強い感情に基づいてフィードバックを返した。本研究ではこれを親和動機とし、実験参加者が社会的な反応を示すことを期待した。(3)に関しては、発話速度は9.0(morae/sec)に設定した。(4)に関しては、実験参加者によって入力された回答に基づいてフィードバックを返した。具体的には、入力された回答と、事

前に単語データベースとして登録された単語を照合し、その結果に基づいてフィードバックを行った。単語データベースには、統制条件と機能低条件で高頻度で回答された単語を予め登録しておき、それらの単語と入力された回答が一致すると、機能低条件と同じように応答した。なお、データベースに登録された単語の数は175語であった。単語データベースに登録されていない単語が回答されると「こちらですね。難しい表現ですね」と応答を返した。

2.3 装置

パーソナルコンピュータを2台、用いた。そのうち1台はHEWLETT-PACKERD社製のノートパソコン(HP ProBook 450 G2)であり、OSはMicrosoft Windows 8であった。実験課題はこのPCで行った。もう1台は教示で差別化するため設置したダミー端末であり、実験には使用しなかった。エージェントへの注視行動を測定するため、眼球運動測定器であるThe Eye Tribe社製のTHE EYE TRIBE TRACKERを用いた。

2.4 実験課題

先行研究[4]の実験課題を改変した。図1に課題画面の例を示す。この課題は2枚の画像から言語で指定された画像を選択する画像選択フェーズと、同じ画像の名称を答える言語入力フェーズで構成されている。画像選択フェーズでは、エージェントが2枚の画像のうちいずれかを言葉で指定し、実験参加者は正しいと思うほうを選択することが要求された。なお、実験参加者が言語同調するかを検討するために、遠回しな言葉も使われた。例えば、「たまねぎ」の画像に対して「球根」という単語が扱われた。実験参加者が画像を選択すると、エージェントが正解か不正解かを応答した。次の言語入力フェーズでは、直前の画像選択フェーズにてエージェントが言及した画像が黄色の枠に囲まれ、実験参加者はその画像の名称を言葉で入力した。その言葉に基づいてエージェントはいずれかの画像を選択した。画像選択フェーズと言語入力フェーズが25回繰り返された。

2.5 印象評価アンケート

実験課題中にエージェントに対してどのような印象を感じていたのかを検討するために、エージェントの

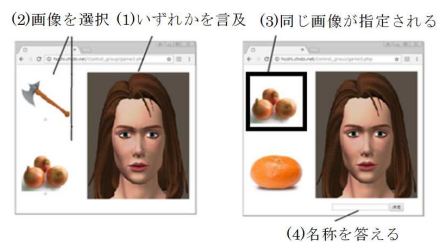


図1 課題画面の例(左: 画像選択フェーズ 右: 言語入力フェーズ)

印象評価アンケートを実施した。印象評価は、Baylorら[23]が用いた、知性と人間らしさを含む項目を参考にコミュニケーションの項目として利用可能な項目を精査し、本研究の指標として利用した。この尺度は、25項目から構成されており、4因子構造(集中できる/知性/人間らしさ/魅力的な)であり、実験参加者は5件法で回答した。なお、本研究では「知性」と「人間らしさ」を分析の対象とした。

2.6 実験システム

WebサーバであるApache(Version 2.4.23)上で動作するシステムとして、PHP(Version 7.0.7)、JavaScript、HTMLを用いて実験課題を作成した。実験参加者の回答はMySQL(Version 5.7.16)によるデータベースに格納された。

2.7 手続き

実験参加者を実験室に入室させ、教示なしの場合、「標準版」のラベルの置かれたPC前に座らせた。教示ありの場合、「高性能」のラベルが置かれたPC前に座らせ、「こちらのエージェントは性能が高く、柔軟に応答することが可能です。」と教示した。その後、実験課題を開始した。画像選択フェーズと言語入力フェーズが25回繰り返されると実験を終了した。実験時間は全体で20分程度であった。

3. 結果

3.1 印象評価

3.1.1 「知性」因子

実験参加者はインタラクションを通して、エージェントに対してどれほど知的に感じていたのかを分析した。各群における、「知性」因子の5件法による得点の平均値を算出したものを図2に示した。

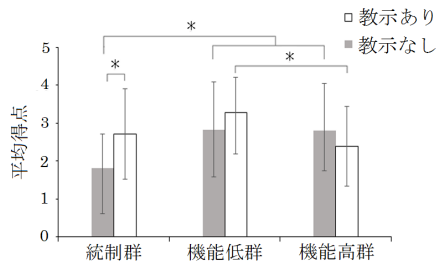


図2 「知性」因子の得点

教示要因と振舞い要因の2要因について、参加者間分散分析を実施したところ、教示と振舞いの交互作用が認められた ($F(2, 84) = 4.5, p < .05, \eta^2 = .09$)。教示の主効果は認められず、振舞いの主効果は認められた ($F(1, 84) = 2.9, p = .10, \eta^2 = .03$; $F(2, 84) = 6.5, p < .005, \eta^2 = .13$)。振舞い要因の単純主効果の検定を実施したところ、統制条件では、教示なし条件よりも教示あり条件のほうが得点が高かった ($F(1, 84) = 8.1, p < .01, \eta^2 = .08$)。動作を伴わないエージェントの場合、知的であるという既有知識があることによって、より知的に感じられていたといえる。機能低条件と機能高条件では、教示の単純主効果が認められなかった ($F(1, 84) = 2.1, p = .15, \eta^2 = .02$; $F(1, 84) = 1.7, p = .19, \eta^2 = .02$)。動作を伴うエージェントの場合、知的であるという既有知識は印象に影響しなかったと考えられる。

続いて、教示要因の単純主効果の検定を実施したところ、教示なし条件においては振舞いの単純主効果が認められた ($F(2, 84) = 6.8, p < .005, \eta^2 = .14$)。Ryan法による多重比較の結果、統制条件よりも機能低条件 ($p < .005$)、統制条件よりも機能高条件において得点が高かった ($p < .005$)。なお、以降の多重比較もRyan法によるものとする。エージェントに対する既有知識がない場合、振舞いに基づいて知性を感じられていたといえる。また、教示なし条件においても振舞いの単純主効果は認められた ($F(2, 84) = 4.2, p < .05, \eta^2 = .09$)。多重比較の結果、機能低条件よりも機能高条件のほうが得点が低かった ($p < .01$)。エージェントに対して知的であるという既有知識があると、一定の動作を伴うよりも、人間の態度に応じて柔軟に振舞うエージェントに対して知的に感じられないといえる。

3.1.2 「人間らしさ」因子

実験参加者はインタラクションを通して、エージェントに対してどれほど人間らしく感じていたのかを分

析した。各群における、「人間らしさ」因子の5件法による得点の平均値を算出したものを図3に示した。

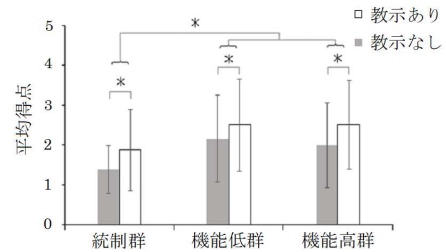


図3 「人間らしさ」因子の得点

教示要因と振舞い要因の2要因について、参加者間分散分析を実施したところ、教示と振舞いの交互作用は認められなかった ($F(2, 84) = .07, p = .92, \eta^2 = .00$)。一方、教示の主効果が認められ ($F(1, 84) = 6.0, p < .05, \eta^2 = .06$)、教示なし条件よりも教示あり条件のほうが得点が高かった。エージェントに対して知的であるという既有知識があると、人間らしく感じられるといえる。また、振舞いの主効果も認められた ($F(2, 84) = 5.8, p < .005, \eta^2 = .12$)。主効果における多重比較の検定の結果、統制条件よりも機能低条件 ($p < .005$)、統制条件よりも機能高条件において得点が高かった ($p < .01$)。既有知識にかかわらず、エージェントが動作を伴うことによって人間らしく感じられていたといえる。

3.2 眼球運動

実験参加者が課題に従事している間の視線データを計測した。図4は、課題中にエージェントを見ていた比率を群ごとに示す。分母を実験の全体時間とし、エージェントを注視していた時間の比率を算出した。

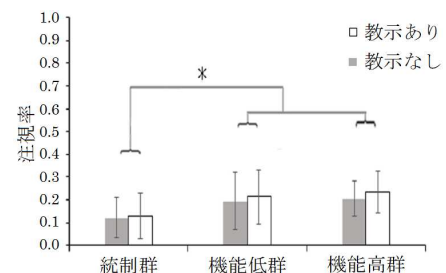


図4 エージェントへの注視行動

分析にあたっては、参加者ごとの比率データに逆正弦変換を行った。教示要因と振舞い要因の2要因について、参加者間分散分析を実施したところ、教示と振舞いの交互作用は認められなかった ($F(2, 84) = .76,$

$p = .92, \eta^2 = .00$). また、教示の主効果は認められず、振舞いの主効果は認められた ($F(1, 84) = .59, p = .44, \eta^2 = .01$; $F(2, 84) = 4.5, p < .05, \eta^2 = .14$). 主効果における多重比較の結果、統制条件よりも機能低条件 ($p < .005$), 統制条件よりも機能高条件において得点が高かった ($p < .005$). 従って、エージェントへの注視行動は、エージェントに対する既有知識の影響を受けず、動作を伴うことによって増加することが明らかになった。

4. 考察

4.1 トップダウン処理の影響

印象評価において、動作を伴わない統制条件では、エージェントが知的であるという既有知識によって知的に人間らしく感じられていた。このことから、エージェントの振舞いから得られる情報が少なければ、既有知識に基づいたトップダウン処理によって印象を形成することが示唆された。

適度に振舞う機能低条件と人間の態度に応じて振舞う機能高条件では、エージェントが知的であるかという既有知識は知的な印象には影響しなかったが、人間らしさには影響していた。これは、知性に関しては、エージェントが動作と音声による発話を伴ったため、実際の振舞いを見て知的かどうかを判断していたことから、相手から得られる情報がある場合、トップダウン処理は弱まることが示唆された。一方、人間らしさに関しては、既有知識の影響があることによって増大したことは、相手への既有知識がコミュニケーションにおける感情特性に影響を与える [2] ことがかわっており、人間らしさは知性以外の要素もあることが示唆される。

4.2 ボトムアップ処理の影響

印象評価において、既有知識を与えない教示なし条件では、エージェントが動作を伴うことによって知的に人間らしく感じられていた。このことから、エージェントに関する既有知識がなければ、実際の振舞いに基づくボトムアップ処理によって印象を形成することが示唆された。しかし、一定の振舞いを表出する機能低条件と人間の態度に応じて振舞う機能高条件では差がみられなかったことは石井・中野 [15] の知見とは異なっており、実験参加者はエージェントの動作は見えていたものの、意図までは感じ取れていなかったことが示唆される。

既有知識を与える教示あり条件では、人間らしさに関しては既有知識を与えない場合と同様に、動作を伴うことによって人間らしく感じられていた。しかし、知性に関しては、知的であるという既有知識があると、人間の態度に応じて振舞う場合は一定の動作よりも、かえって知的に感じられなかった。これは、複雑な動きが、知的であるという既有知識に適合しなかったため、適応ギャップ [11] が表出したと考えられる。本実験のシステムは完璧なものではなく、エージェントが誤ったフィードバックをすることがあったことが原因であると考えられる。よって、今後はより精度の高いシステムを開発する必要がある。

眼球運動に関しては、エージェントが動作を伴うことによって注視行動が増えた。これは、既有知識は影響せず、相手の振舞いに基づいたボトムアップ処理のみが影響していたと考えられる。

4.3 2種類の処理の相互作用

以上から、印象評価に関しては、トップダウン処理とボトムアップ処理の利用が異なっていたことから、仮説 1-c が部分的に支持されたといえる。また、注視行動に関しては、エージェントの振舞いのみに影響を受けていたことから、仮説 2-b が支持されたといえる。

本研究の結果で興味深い点は、知性と人間らしさの印象が一貫しなかった点である。知性に関しては、エージェントが知的であるという既有知識がある場合において、適応ギャップが表出し、知的に感じられなかった。一方で、人間らしさに関しては、知的であるという既有知識があることによって、振舞いにかかわらず人間らしく感じられた。エージェントを対人認知の対象として考えたとき、一般に人間は主に能力面を中心的な問題として捉えることが明らかになっている [19]。だが、能力以外の点に関する対人認知が適切に行われるならば、この適応ギャップを解決するための端緒となることが議論されている [24]。本実験のエージェントは、実験参加者の感情に応じた感情表出を行っていたことを中心に、能力面以外でも人間らしさを形成していた。このことが、人間らしさにおいては適応ギャップを回避していたと考えられる。このことから、エージェントに人間らしさを実装する場合、能力面以外の要素（感情など）をデザインすることが、適応ギャップを抑制する上ではひとつの解決策になることが示唆される。

4.4 全体のまとめ

本研究では、エージェントとの会話場面において、どのような認知情報処理がエージェントの印象とエージェントへの注視行動にどう影響するのかを検討した。特にここでは、(1) 既有知識に基づくトップダウン処理と、(2) 相手の振舞いに基づくボトムアップ処理の2つに着目した。仮説として、エージェントに関する既有知識はトップダウン処理として印象形成に影響を及ぼすが、相手の振舞いから得られる情報量に応じてその影響力は弱まり、ボトムアップ処理が優勢になると予想した。また、注視行動は相手の動きに影響されるため、ボトムアップ処理のみの影響があると予想した。この仮説を検証するため、言語選択課題を用いて、(1) エージェントが知的であるかどうかという既有知識の要因と、エージェントの知的な振舞いの2×3の参加者間要因の実験を行った。その結果、(a) エージェントの振舞いから得られる情報量が少なければ既有知識に基づくトップダウン処理によって印象形成を行っていたが、(b) 相手の振舞いが顕著になるにつれてボトムアップ処理によって印象形成を行うことがわかった。

謝辞

本研究は JSPS 科研費 16K00219 を受けたものです。ここに記して謝意を表す。

参考文献

- [1] G.V. Bodenhausen and R.S. Wyer, "Effects of stereotypes in decision making and information-processing strategies," *Journal of Personality and Social Psychology*, vol.48, no.2, p.267, 1985.
- [2] 林 勇吾, 三輪和久, "人間とエージェントが混在する状況におけるコミュニケーションの認知的・感情的心理特性," *ヒューマンインタフェース学会論文誌*, vol.10, no.4, pp.445-456, 2008.
- [3] S. Wu and B. Keysar, "The effect of information overlap on communication effectiveness," *Cognitive Science*, vol.31, no.1, pp.169-181, 2007.
- [4] J. Pearson, J. Hu, H.P. Branigan, M.J. Pickering, and C.I. Nass, "Adaptive language behavior in HCI: how expectations and beliefs about a system affect users' word choice," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp.1177-1180, 2006.
- [5] C.L. Kleinke, "The role of gaze in impression formation," *British Journal of Clinical Psychology*, vol.14, no.1, pp.19-25, 1975.
- [6] M. Argyle, L. Lefebvre, and M. Cook, "The meaning of five patterns of gaze," *European Journal of Social Psychology*, vol.4, no.2, pp.125-136, 1974.
- [7] P.C. Ellsworth and J.M. Carlsmith, "Effects of eye contact and verbal content on affective response to a dyadic interaction," *Journal of Personality and Social Psychology*, vol.10, no.1, p.15, 1968.
- [8] P.H. Waxer, "Nonverbal cues for anxiety: An examination of emotional leakage," *Journal of Abnormal Psychology*, vol.86, no.3, p.306, 1977.
- [9] 神田崇行, 石黒 浩, 石田 亨, "人間-ロボット間相互作用にかかわる心理学的評価," *日本ロボット学会誌*, vol.19, no.3, pp.362-371, 2001.
- [10] D.A. Norman, *誰のためのデザイン-認知科学者のデザイン原論-*, 新曜社, 1990.
- [11] 小松孝徳, 山田誠二, "適応ギャップがユーザのエージェントに対する印象変化に与える影響," *人工知能学会論文誌*, vol.24, no.2, pp.232-240, 2009.
- [12] 内田照久, "音声の発話速度の制御がピッチ感及び話者の性格印象に与える影響," *日本音響学会誌*, vol.56, no.6, pp.396-405, 2000.
- [13] M. Argyle and M. Cook, *Gaze and mutual gaze*, Cambridge U Press, 1976.
- [14] Y.I. Nakano, G. Reinstein, T. Stocky, and J. Cassell, "Towards a model of face-to-face grounding," *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1*, pp.553-561, 2003.
- [15] 石井 亮, 中野有紀子, "ユーザの注視行動に基づく会話参加態度の推定—会話エージェントにおける適応的会話制御に向けて," *情報処理学会論文誌*, vol.49, no.12, pp.3835-3846, 2008.
- [16] 中野有紀子, 岡 兼司, 佐藤洋一, 西田豊明, "ユーザの視線に気づく会話エージェント-アテンションの知覚と制御を利用した会話の円滑化," *人工知能学会全国大会論文集 2005 年度人工知能学会全国大会 (第 19 回) 論文集*, p.221, 2005.
- [17] 小孫康平, "瞬目の多少が人の印象形成に及ぼす影響," *日本教育工学会論文誌*, vol.30, no.30, pp.1-4, 2006.
- [18] D. Davis, "Implications for interaction versus effectance as mediators of the similarity-attraction relationship," *Journal of Experimental Social Psychology*, vol.17, no.1, pp.96-117, 1981.
- [19] 竹内勇剛, 片桐恭弘, "ユーザの社会性に基づくエージェントに対する同調反応の誘発," *情報処理学会論文誌*, vol.41, no.5, pp.1257-1266, 2000.
- [20] "Web 便利ノート," <https://note.cman.jp/other/voice>, 参照 Aug. 1, 2016.
- [21] "Javascript library for precise tracking of facial features via constrained local models," <https://github.com/auduno/clmtrackr>, 参照 Aug. 1, 2016.
- [22] D. Davis, "Unmasking the face: A guide to recognizing emotions from facial cues," 1975.
- [23] A. Baylor and J. Ryu, "The API (Agent Persona Instrument) for assessing pedagogical agent persona," *EdMedia: World Conference on Educational Media and Technology*, pp.448-451, 2003.
- [24] 鈴木紀子, 竹内勇剛, 石井和夫, 岡田美智男, "状況に引き出された発話による対話の形成とその心理的評価," *情報処理学会論文誌*, vol.40, no.4, pp.1453-1463, 1999.