

身体操作を実装した対話エージェントとの持続的インタラクション評価

Development and analysis of adaptors in a continuing interaction with a conversational agent

東野寛志¹ 神田智子^{1,2}

Hiroshi Higashino¹, Tomoko Koda^{1,2}

¹大阪工業大学大学院情報科学研究科情報科学専攻

¹Graduate School and Faculty of Information Science and Technology, Osaka Institute of Technology

²大阪工業大学情報科学部情報メディア学科

²Department of Media Science, Faculty of Information Science and Technology, Osaka Institute of Technology

Abstract: Self-adaptors are bodily behaviors that often involve self-touch. Our evaluation of the interaction between an agent with self-adaptors and without indicated that there is a dichotomy on the evaluation of interaction between users with high social skills and those with low skills. People with high social skills feel more friendliness toward an agent that exhibits self-adaptors than those with low social skills. The result suggests the need to tailor non-verbal behavior of virtual agents according to user's social skills.

1. はじめに

対面で人とのインタラクションを行う会話エージェントが一般ユーザに普及し始めており、盛んに研究が行われている[1]。会話エージェントは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。ノンバーバルコミュニケーションは身体的特徴や対人距離、周辺言語、身体動作などに分類される。Ekman は身体動作を標識、例示子、情感表示、調整子、適応子に分類している[2]。本研究では身体動作のうち適応子と呼ばれる動作群の身体操作に着目する。身体操作は鼻の穴をほじる、頭をかくのような身体のある部分で他の身体のある部分に加える操作のことであり、人前でタブーとされる動作が多く、さらにメッセージ性が低く、対話内容との関連性が低いとされる[3]。そのため HAI 研究において、表情や視線などメッセージ性の強いノンバーバルコミュニケーションの研究と比べると活発に研究が行われておらず、身体操作がインタラクション上にもたらす肯定的な効果が示されていない。例えば、Neff らの研究では、頭や顔を搔く、身体を触る、手をこすり合わせるなどの身体

操作を実装した対話エージェントは、見かけの精神的安定度が低下することが報告されている[4]。

しかし、人同士の親しい間柄でのコミュニケーションにおいて、身体操作は頻繁に行われている。このことから、人同士のコミュニケーションにおいて、身体操作は親近感を抱いた対話相手に対して多く行う動作であることと、対話相手が身体操作を行うことで、親近感を抱くことが考えられる。ユーザと対話を行う対話エージェントの身体動作に身体操作があることで、ユーザの対話エージェントへの親近感を誘発することが可能ではないかと考えた。

身体操作はメッセージ性が低く、対話内容との関連性が低いことから、対話中では見逃してしまいやすい動作であると考えられる。対人関係を円滑にする個人特性として、社会的スキルがある。社会的スキルとは「対人関係を円滑にはこぶために役立つスキル(技能)」と定義されている[5]。人同士のコミュニケーションにおいて、社会的スキルの高い人は、対話相手のノンバーバルな行動を読み取り、コミュニケーションに活かすことができると考えられる。さらに、社会的スキルが高い人は対話相手とのコミュニケーションをより豊かなものにするため、ノン

バーバルな行動を多く用いる傾向があると考えられる。この社会的スキルの特性に着目し、人対エージェントのコミュニケーションにおいても同じことが言えるのではないかと考えた。

Reevesらは、メディアに対する人の反応は、社会的かつ自然であるとしている[6]。また、メディアと性格に関する実験で、ユーザは自分と同じ性格のコンピュータを好む傾向にあることを示した[6]。このことから、ユーザは自分と同じ性格のエージェントを好むのではないかと考えられる。社会的スキルの特性から、社会的スキルが高い人は、身体操作を行う対話エージェントに対して、自分と同じ性格であると考えたと推測した。そこで本研究では、「社会的スキルの低い人に比べて、社会的スキルが高い人は身体操作を実装した対話エージェントにより親近感を抱く」と仮説を立て、検証を行った。

エージェントとユーザとのインタラクションに関する研究は数多く行われているが、多くが一過性のインタラクション評価であり、エージェントとユーザとの持続的なインタラクション評価についての研究は数少ない。その中でも代表的なのが Bickmoreらの Relational Agents 関連の研究がある。Bickmoreらによれば、ユーザとエージェントとの持続的なインタラクションには信頼関係を築くことが重要であると述べている[7]。本研究では、インフォーマルコミュニケーションにおける身体操作が、複数回のインタラクションを経て形成されるのを鑑み、一度の実験で印象評価を行うのではなく、実験を複数回、継続的に行い印象評価を行う事で、エージェントとユーザとの持続的なインタラクション評価を行うこととした。

2. 対話エージェント

対話エージェントに実装する身体操作は先行研究[8]で行われた、友人同士の対話における身体操作のビデオ分析結果を参考に「髪を触る」、「顔を触る」、「鼻を触る」の3種類とした。エージェントは Poser7 を用いて作成し、3種類の身体操作を行うアニメーションを作成した。図1に身体操作「髪を触る」「鼻を触る」を行うエージェントを示す。また、身体操作以外の動作として、ジェスチャー「首をかしげる」「手を胸にあてる」を行うエージェントのアニメーションを作成した。これらのジェスチャーは対話内容に沿って適切なタイミングで行う。

ユーザとエージェントの対話方式は、エージェントの質問に対する回答を参加者が選択肢の中から選択して回答する、選択式の対話とし、エージェントの対話システムを Microsoft Visual Studio 2008 を使用し、C++で開発を行った。エージェントの音声には

音声合成パッケージ AITalk 声の職人[9]を使用した。エージェント対話システムの概念図を図2に示す。ユーザとエージェントの対話内容は、事前に対話シナリオを作成する。その対話シナリオに沿ってエージェントのアニメーションを映像として作成し、ユーザの回答内容に沿った映像をつなぎ合わせて、疑似的にユーザとの対話を実現する。対話システムには2つの状態があり、交互に遷移することでユーザとの対話を実現する。1つは発話状態で、エージェントがユーザに対して発話している映像を再生する状態である。もう1つはユーザ選択待機状態でエージェントの質問に対して、ユーザが選択肢から回答を選択し、入力する状態である。入力された回答に対して、会話シナリオに沿い対応する映像を発話状態として再生する。

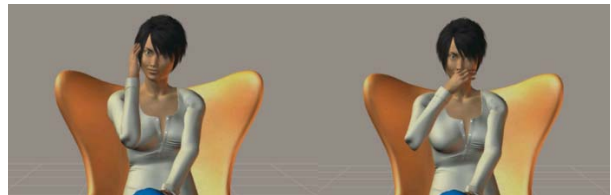


図1. 身体操作「髪を触る」(左), 「鼻を触る」(右)を行うエージェント

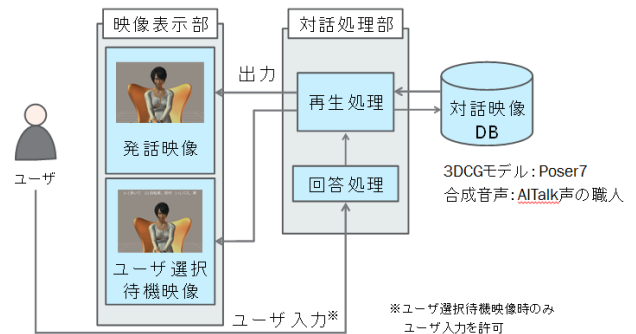


図2. エージェント対話システムの概念図

3. 実験方法

実験参加者は24名の男子大学生と大学院生である。KiSS-18[10]を用いて社会的スキルを測定し、事前の実験参加者を社会的スキル高群と社会的スキル低群に分類しておいた。日本人成人男性、女性の平均得点がそれぞれ、61.82, 60.10であるため[10]、これを参考に社会的スキル尺度得点の63点以上を社会的スキル高群(11名)、58点以下を社会的スキル低群(13名)とした。

社会スキル高群と社会的スキル低群の実験参加者は、身体操作を行う対話エージェント(社会スキル高群7名、社会スキル低群7名)、身体操作を行わない対話エージェント(社会スキル高群4名、社会スキル

低群6名)のいずれかのエージェントのみと計5回の対話を行う。エージェントとの対話は1日1回とし、対話するエージェントの種類(身体操作の有無)は全実験回で同じとする。対話エージェントの種類の違いは、身体操作を行うか行わないかだけの違いで、エージェントの外見、音声、ジェスチャーのタイミングと回数、対話内容ともに同じである。また、対話内容は全実験回数で異なるよう、5回分の対話シナリオを用意した。実験回数における対話シナリオの順番は、エージェントの種類に関係なく同じである。実験の条件は社会的スキル(社会的スキル高群、社会的スキル低群)とエージェントの種類(身体操作を行うエージェント、身体操作を行わないエージェント)と実験回数(1回目、2回目、3回目、4回目、5回目)である。

各対話回の最後に、6段階評定のSD法によりエージェントの印象評価を行う。印象評価項目は対人認知の特性形容詞尺度[11]の20項目と独自の3項目(エージェントの「人間らしさ」「わずらわしさ」「自然さ」)を合わせた23項目の形容詞対を用いた。また、5回目の対話実験終了後に、実験終了後アンケート調査を実施し、エージェント動作や合成音声の自然さ等の主観的評価を行った。

4. 結果

4.1 親近性因子の分析

実験の結果、得られた対話エージェントの印象評価値を用い、因子分析を行った結果、「親近性因子」「慎重性因子」「寛容性因子」の3種類の因子が抽出された。主因子法を用い因子分析を行った結果を表1に示す。人が対話エージェントを対人的に認知し、印象評価する際は、これら3つの因子が大きく影響していることがわかった。因子分析により抽出された各因子と関係が強い(因子負荷量の大きい)項目の測定値を単純に合算し、その合成値を因子の測定尺度値とする。因子分析で抽出された親近性因子、慎重性因子、寛容性因子それぞれの測定尺度値を用い、社会的スキル要因(高群、低群)、身体操作要因(あり、なし)、実験回数要因(1回目、5回目)の3要因の分散分析を行った。

親近性因子では、社会的スキル要因と身体操作要因と実験回数要因の2次の交互作用($p<0.01$)で有意差がみられた。身体操作要因と実験回数要因の各水準の組み合わせにおける社会的スキル要因の多重比較を行った結果を図3に示す。身体操作要因「あり」と実験回数要因「1回目」の組み合わせと、身体操作要因「あり」と実験回数要因「5回目」の組み合わせにおいて、社会的スキル要因に有意差($p<0.05$)

が見られた(ともに高群>低群)。社会的スキル高群は社会的スキル低群に比べ、身体操作を行う対話エージェントに対し、実験回数が1回目、5回目共に、有意に親近性が高いことが示された。

次に、社会的スキル要因と実験回数要因の各水準の組み合わせにおける身体操作要因の多重比較を行った結果を図4に示す。社会的スキル要因と実験回数要因のどの水準の組み合わせにおいても、身体操作要因の「あり」と「なし」に有意差が見られなかった。

社会的スキル要因と身体操作要因の各水準の組み合わせにおける実験回数要因の多重比較を行った結果を図5に示す。社会的スキル要因「高群」と身体操作要因「あり」の組み合わせと、社会的スキル要因「低群」と身体操作要因「なし」の組み合わせにおいて、実験回数要因「1回目」と「5回目」の間に有意差($p<0.05$)が見られた。社会的スキル高群は身体操作を行う対話エージェントに対し、実験回数1回目に比べ、実験回数5回目の方が有意に親近性因子に対する評価が高いことが示された。また、社会的スキル低群は身体操作を行わない対話エージェントに対し、実験回数1回目に比べ、実験回数5回目の方が有意に親近性因子に対する評価が高いことが示された。

慎重性因子の測定尺度値を用い、社会的スキル要因、身体操作要因、実験回数要因の3要因の分散分析を行った。その結果、どの要因の主効果、交互作用においても有意差は見られなかった。

表1. 因子分析結果(プロマックス回転後の因子パターン)

評価項目	因子1	因子2	因子3
人間らしくない - 人間らしい	1.009	-0.098	-2.05
近づきたい - ひとつっこい	.930	-3.09	.118
不自然な - 自然な	.906	.006	-0.30
感じのわるい - 感じのよい	.889	.050	.002
非社交的な - 社交的な	.856	.075	-2.17
無気力な - 意欲的な	.784	.020	-0.44
沈んだ - うきうきした	.784	-1.85	.098
にくらしい - かわいらしい	.775	.069	.135
人のわるい - 人のよい	.736	.219	-0.07
消極的な - 積極的な	.700	-1.40	-1.25
親しみにくい - 親しみやすい	.681	-1.45	.333
不親切な - 親切な	.681	.205	.076
無分別な - 分別のある	.509	.204	-0.89
責任感のない - 責任感のある	.488	.485	.008
軽率な - 慎重な	-.181	.751	-0.69
軽薄な - 重厚な	-.148	.645	.188
なまいきな - なまいきでない	.403	.502	.012
わずらわしい - おとなしい	.069	.434	.155
短気な - 気長な	-.211	.143	.863
心のせまい - 心のひろい	.297	.001	.527
卑屈な - 堂々とした	.162	-0.14	.177
自信のない - 自信のある	.238	-0.006	.096
恥ずかしがりの - 恥知らずの	-.148	-.242	.062

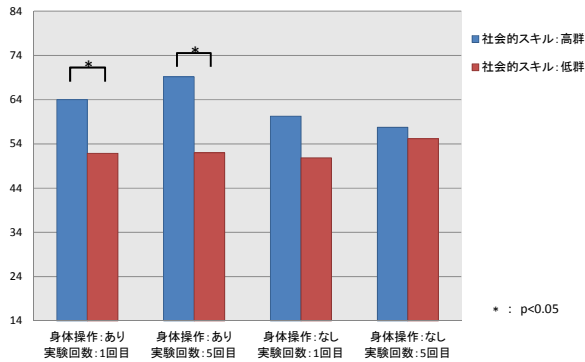


図 3. 身体操作×実験回数における社会的スキル要因別の多重比較結果(親近性因子測定尺度値)

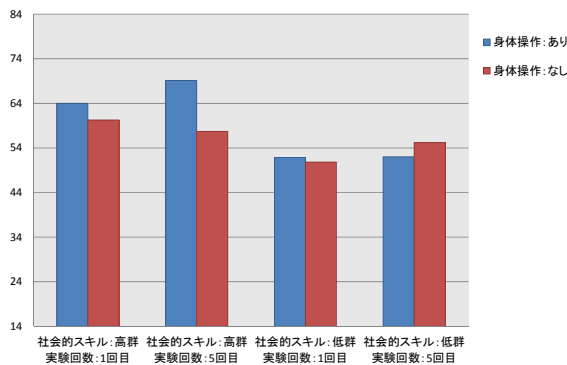


図 4. 社会的スキル×実験回数における身体操作要因別の多重比較結果 (親近性因子測定尺度値)

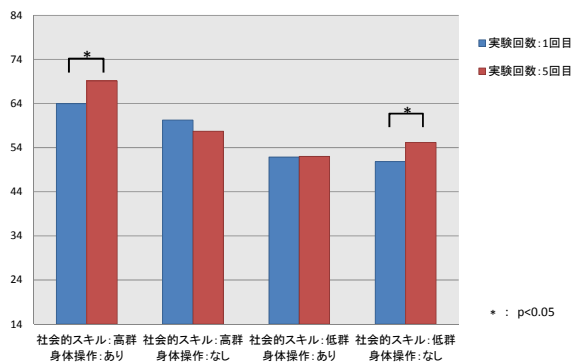


図 5. 社会的スキル×身体操作における実験回数要因別の多重比較結果(親近性因子測定尺度値)

4.2 寛容性因子の分析

寛容性因子の測定尺度値を用い、社会的スキル要因、身体操作要因、実験回数要因の3要因の分散分析を行った。その結果、寛容性因子では、社会的スキル要因と実験回数要因の1次の交互作用で有意差($p < 0.05$)がみられた。実験回数要因の各水準における社会的スキル要因の多重比較を行った結果を図6に示す。実験回数要因「1回目」において、社会的ス

キル要因に有意差($p < 0.05$)がみられた(高群 > 低群)。実験回数が1回目の場合、社会的スキル低群に比べて社会的スキル高群は、対話エージェントに対して有意に寛容的であると評価することが示された。

社会的スキル要因の各水準における実験回数要因の多重比較を行った結果を図7に示す。社会的スキル要因「低群」において、実験回数要因に有意差($p < 0.05$)がみられた(1回目 < 5回目)。社会的スキル低群は対話エージェントに対して実験回数1回目に比べ、実験回数5回目の方が有意に寛容的と評価することが分かった。

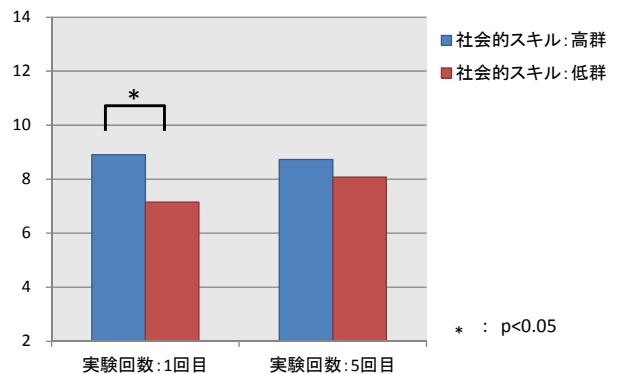


図 6. 実験回数要因の各水準における社会的スキル要因別の多重比較結果(寛容性因子の測定尺度値)

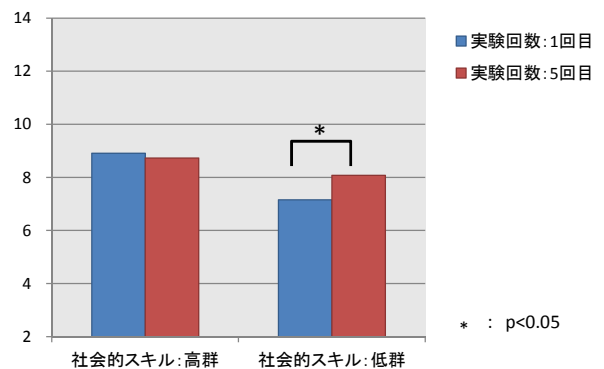


図 7. 社会的スキル要因の各水準における実験回数要因別の多重比較結果(寛容性因子の測定尺度値)

4.3 実験終了後アンケートの分析

実験終了後アンケートの評価値(1:低-8:高の8段階評価)を用い、社会的スキル要因、身体操作要因の2要因分散分析を行った。「エージェントの動作が気になりましたか」について、社会的スキル要因と身体操作要因の交互作用で有意差($p < 0.05$)がみられた。社会的スキル要因の各水準における身体操作要因の多重比較を行った結果、社会的スキル「低群」において、身体操作要因に有意差($p < 0.05$)がみられた(あり:5.43 > なし:3.83)。社会的スキル低群は身体操作を

行わない対話エージェントに比べ、身体操作を行う対話エージェントの方が、有意に動作が気になると評価したことが示された。

「エージェントの音声は聞き取りやすかったですか」について社会的スキル要因の主効果で有意差 ($p < 0.05$) がみられた。社会的スキル低群(4.69)に比べ、社会的スキル高群(6.18)は有意にエージェントの音声が聞き取りやすかったと評価したことが示された。

5. 考察

5.1 親近性因子の分析結果の考察

図 3 より、社会的スキル高群は社会的スキル低群に比べ、身体操作を行う対話エージェントに対し、実験回数が 1 回目、5 回目共に、有意に親近性が高いことが示された。このことから、実験回数にかかわらず、身体操作を行う対話エージェントに対して社会的スキル高群は、社会的スキル低群に比べて有意に高い親近感を抱くといえる。また、身体操作要因と実験回数要因の各水準の組み合わせにおける社会的スキル要因「低群」の値にあまり差がないことから、身体操作を行う対話エージェントに対して、社会スキル低群が親近性を欠いたのではなく、社会的スキル高群が、身体操作を行う対話エージェントに対して、強く親近性を感じたと考えられる。これは仮説の「社会的スキルの低い人に比べて、社会的スキルの高い人は身体操作を実装した対話エージェントにより親近感を抱く」を支持する結果である。

図 4 より、社会的スキル要因と実験回数要因の各水準の組み合わせにおける身体操作要因の多重比較を行った結果、社会的スキル要因と実験回数要因のどの水準の組み合わせにおいても、身体操作要因の「あり」と「なし」に有意差が見られなかった。このことから、身体操作要因だけでは、親近性に差がないことがわかる。しかし、有意差こそなかったものの、社会的スキル要因「高群」では身体操作要因「あり」と「なし」の差が大きく、社会的スキル「低群」では身体操作要因「あり」と「なし」の差が小さいことがわかる。このことから、身体操作を行う対話エージェントに対して、社会的スキル低群が親近性を欠いたのではなく、社会的スキル高群が、身体操作を行う対話エージェントに対して、強く親近性を感じたという考えを支持する結果であるといえる。

図 5 より、社会的スキル高群は身体操作を行う対話エージェントに対し、実験回数 1 回目に比べ、実験回数 5 回目の方が有意に親近性因子に対する評価が高いことが示され、社会的スキル低群は身体操作を行わない対話エージェントに対し、実験回数 1 回

目に比べ、実験回数 5 回目の方が有意に親近性因子に対する評価が高いことが示された。このことから、対話エージェントとの持続的インタラクションにおいて、社会的スキル高群は対話エージェントが身体操作を行うことで、親近性を増していくのに対し、社会的スキル低群は対話エージェントが身体操作を行わないことで親近性を増していくといえる。これも仮説の「社会的スキルの低い人に比べて、社会的スキルの高い人は身体操作を実装した対話エージェントにより親近感を抱く」を支持する結果である。また、対話エージェントとの持続的インタラクションにおいて、社会的スキル高群は身体操作を行う対話エージェントの親近性が向上していくのに対し、社会的スキル低群は、身体操作を行わない対話エージェントの親近性が向上していくことから、社会的スキル高群と社会スキル低群とで対話エージェントの身体操作についての親近性評価が二分化していることが考えられる。

これらの結果から、仮説の「社会的スキルの低い人に比べて、社会的スキルの高い人は身体操作を実装した対話エージェントにより親近感を抱く」は支持されたといえる。また、身体操作を対話エージェントに実装する際には、ユーザの社会的スキルの高低に合わせて、対話エージェントを開発する必要があること、エージェントとの持続的インタラクションにおいては、社会的スキルの高低と身体操作の実装の有無の組み合わせに考慮することで、ユーザのエージェントに対する親近性を持続的に向上させることができることが示唆されたと考えられる。

5.2 寛容性因子の分析結果の考察

図 6 より、実験回数が 1 回目の場合、社会的スキル低群に比べて社会的スキル高群は、対話エージェントに対して有意に寛容的と評価することが示された。図 7 より、社会的スキル低群は対話エージェントに対して実験回数 1 回目に比べ、実験回数 5 回目の方が有意に寛容的と評価することが分かった。

これらの結果を踏まえると、社会的スキル低群は社会的スキル高群に比べ、対話エージェントを寛容的だと評価しないが、対話エージェントとの持続的なインタラクションによって社会的スキル低群は、寛容性の評価を上げていくと考えられる。社会的スキル高群は対話エージェントとの持続的インタラクションにおいて、対話エージェントの寛容性の評価が上がることはないが、初回インタラクション時から対話エージェントを社会的スキル低群と比べ、より寛容的だと評価していることが考えられる。

5.3 実験終了後アンケートの考察

実験終了後アンケートの質問項目である「エージェントの動作が気になりましたか」について、社会的スキル低群は身体操作を行わない対話エージェントに比べ、身体操作を行う対話エージェントの方が、有意に動作が気になると評価したことが示された。この結果は、親近性因子において、社会スキル低群が身体操作を行う対話エージェントに対する評価が低かったことと関連した結果であると考えられる。エージェントの動作が気になったということが親近性因子の評価にネガティブに作用したことが考えられ、社会的スキル低群は「動作が気になった」が「動作が気に食わない」と同値の評価であることが考えられる。社会的スキル要因「高群」では身体操作要因「あり」と「なし」の間に差がなかったことから、社会的スキル低群は対話エージェントの身体操作について、過敏に反応し、ネガティブな評価を持つことが示唆された。

「エージェントの音声は聞き取りやすかったですか」について、社会的スキル低群に比べ、社会的スキル高群は有意にエージェントの音声の聞き取りやすかったと評価したことが示された。対話エージェントの音声は、社会的スキル高群と低群とでまったく同じものを使用しており、社会的スキル高群は対話エージェントの評価に関し肯定的であるが、低群は対話エージェントの評価に関し否定的である可能性があることを示唆している。

6. 結論

本研究では、「社会的スキルの低い人に比べて、社会的スキルの高い人は身体操作を実装した対話エージェントにより親近感を抱く」と仮説を立て、実験を行いその検証を行った。実験から得られた印象評価の結果を用い、因子分析を行った結果、「親近性因子」「慎重性因子」「寛容性因子」の3種類の因子が抽出された。人が対話エージェントを对人的に認知し、印象評価する際は、これら3つの因子が大きく影響していることがわかった。対話エージェントの開発には、親近性、慎重性、寛容性に着目することでユーザに受け入れられやすい対話エージェントの開発が可能であると考えられる。

仮説の「社会的スキルの低い人に比べて、社会的スキルの高い人は身体操作を実装した対話エージェントにより親近感を抱く」は支持された。このことから、身体操作を対話エージェントに実装する際には、ユーザの社会的スキルの高低に合わせて、開発する必要がある。また、持続的なインタラクションにおいて、身体操作を行う対話エージェントの親近性評価は、社会的スキル高群と社会的スキル低群とで二分化されることが示唆された。エージェントと

の持続的インタラクションにおいては、社会的スキルの高低と身体操作の実装の有無の組み合わせに考慮することで、ユーザのエージェントに対する親近性を持続的に向上させることができると示唆された。

今後は、ユーザの社会的スキル高低別に、対話エージェントが身体操作を行う適切なタイミングや回数をより緻密に調査することで、ユーザの社会スキルが高い人、低い人に合わせて効率よくユーザの親近性を引き出すことができると考えられる。また、対話エージェントとの持続的なインタラクションの中で、ユーザ動作から社会的スキルを測定することで、身体操作の回数、頻度を自動制御し、ユーザの親近感を引き出すことができるエージェントの開発が可能で、より質の高いエージェントインタラクションが期待できる。

謝辞

本研究の一部はJSPS 科研費基盤(C)23500266の助成を受けたものである。

参考文献

- [1] 山田 誠二, 人とロボットの<間>をデザインする, pp. 3-22. 東京電機大学出版局 (2007)
- [2] Ekman P, Three classes of nonverbal behavior, *Aspects of Nonverbal Communication*, Swets and Zeitlinger (1980)
- [3] 黒川 隆夫. ノンバーバルインタフェース, pp.1-68, オーム社 (1994)
- [4] Neff, M., Toothman, N., Bowmani, R., Fox Tree, J. E., Walker, M. Don't Scratch! Self-adaptors Reflect Emotional Stability. In: Vilhjalmsson, H. H. et al. (Eds.): IVA 2011, LNAI 6895, pp. 398-411, Springer-Verlag (2011)
- [5] 菊池章夫, 思いやりを科学する, 川島書店 (1988)
- [6] Reeves, B. and Nass, C. *Media Equation : How People Treat Computers, Television and New Media like Real People and Place*, Univ. of Chicago Press (1996)
- [7] Bickmore, T. and Cassell, J., *Relational Agents: A Model and Implementation of Building User Trust*, In: Proc. of CHI 2001, pp. 396-403 (2001)
- [8] 東野寛志, 神田智子. 身体操作を実装した仮想エージェントとの持続的インタラクション評価. HAI シンポジウム 2010 (2010)
- [9] AITalk 声の職人, 株式会社エーアイ <http://www.ai-j.jp/voice> (2012年10月4日閲覧)
- [10] 菊池章夫, 社会的スキルを測る: KiSS-18 ハンドブック. 川島書店 (2007)
- [11] 林文俊, 対人認知構造における個人差の測定(VIII)―認知者の自己概念および欲求との関連について, 実験社会心理学研究, 22, pp. 1-9 (1982)