

非言語応答の半自律生成機能を持つ社会的複数 CG アバター対話システムの評価

○上原悠輝（大阪大学） 酒井和紀（大阪大学） 吉川雄一郎（大阪大学）
石黒浩（大阪大学）

1. 緒言

近年、遠隔地間でのコミュニケーションツールとして、Zoom、Skype などの無料でビデオ対話を行えるものが普及して、容易にビデオ対話を行える環境が整ってきた。また技術の進歩により、複数人での遠隔コミュニケーションもより手軽に行うこともできるようになってきた。しかし、ビデオ対話システムには問題点が2つ挙げられる。1つ目は、複数人のビデオ対話において、対話者が誰に視線を向けているかわからない点である。ビデオ対話では、通常、対話者はカメラのほうを向いているので正面を見ていることになる。そのため誰が誰に視線を向けているのかわからないといった、非言語情報の伝達がうまくいかず、その結果対話が円滑に進まないことが多々ある。実際、複数人ビデオ対話において、顔映像による視線の不一致は対話参加者らのモチベーションを低下させることが報告されている [1]。2つ目に、ビデオ対話において、自身のふるまいが常時見られていると感じることで対話時に不安を感じる可能性がある点である。ビデオ対話は自身を含む全員が映像として映っており、常に自分の映像が複数人の相手に見られている。自分が話していないときにも他の対話者と同じように強調されてしまっており、特に対話に集中しづらい状況にある人は自身の振る舞いがほかの対話者に伝わることに不安感を抱く可能性がある [2]。これらの問題を解決するために、対話者として“自身の代理エージェント”となるアバターを用い仮想空間上で複数人で対話するシステムを導入する。そしてそのアバターに社会的に礼儀正しく振舞わせること（以後、社会的装飾と呼ぶ）を考える。これは、例えば対面での対話において、話し手は対話相手に自身の話に関心を持ってもらおうとふるまい、聴き手は対話相手に関心をもってふるまうといった、我々が対話相手に対し行う非言語応答である。これまで、社会的装飾が施されたアバターが人間と対話する研究はあったが [3]、アバター対話における社会的装飾の効果は不明である。本研究では複数アバター対話システムを用い、それに社会的装飾を半自律に生成し動作させることで、ビデオ対話システムと比べ、常に対話に集中することが難しい人が与える被傾聴感が改善する、対話に集中しづらい人を社会的に礼儀正しいと感じる、対話に集中しづらい人の対話不安が改善される、という効果があるという仮説を立てた。本研究では社会的装飾機能を持つ複数アバター対話システムとビデオ対話システムの2条件を用いて被験者に3人1組で対話させ、心理学的実験により評価を行った。本稿では、本研究で導入した複数アバター対話システム及びそのシステムに導入した社会的装飾機能について説明した後、提案システムを使用することによる効果を検証するためにに行った実験とその報告を行う。



図1 半自律社会的 CG アバタールーム「CommUTalk」

2. 提案システム

2.1 半自律社会的 CG アバタールーム「CommUTalk」

提案する機能を実装するプラットフォームとして、アバター対話システムには CommUTalk を使用した。CommUTalk は大阪大学、ヴィストン株式会社が開発した人型対話ロボット CommU の 3DCG をアバターとして使用しており、システムはブラウザ上で動作する。ブラウザ画面に3体のロボットが表示されており、ロボットは全員同じ仮想空間上にいる。図1中の手前のロボットが自身のアバターであり自身のアバターは半透明表示されている。各参加者は音声で対話に参加する。3人までのアバター対話が可能である。

2.2 社会的装飾機能の実装

CommUTalk に以下のような社会的装飾機能を実装した。1つ目は発話によって他2体のアバターが発話者のアバターのほうに視線を向ける機能である。これは対話者の発声が誰から発せられたものかを判断し、他2体のアバターを指定して動作させる仕組みとなっている。2つ目は対話者が発話していることが分かるよう、発話者のアバターの腕が動く機能である。これは発話者の発話音声から抽出されたフォルマントをもとに推定された母音に応じてアバターの腕の動きを変化させている。3つ目は発話終了時、聴き手2人が一定の確率で5種類の“傾聴動作”を行う機能である。これはそれぞれ動作が異なる4種類の頷きの動作（小さく頷く、大きく頷く、ゆっくり頷く、2回頷く）と何もしない（頷かない）動作をランダムで行う仕組みになっている。

3. 実験

社会的装飾を施した複数アバター対話システムを使用し対話した際の効果を検証するための被験者実験を行った。本実験では、社会的装飾機能を強調するために、対話に集中している/集中しづらい状況にある人で対話する状況を作る。仮説は3点ある。対話に集中しづらい状況において、社会的装飾機能を施した複数

アバター対話システムを用いることで、ビデオ対話システムよりも、対話に集中しづらい状況にある人からの傾聴感が改善する、対話に集中しづらい人でも高い社会的スキルを持っているように感じさせる、対話に集中しづらい人の対話不安感が減少する、である。そこで、社会的装飾を行った複数アバター対話システム条件とビデオ対話システム条件を用意し、被験者内比較を行った。

3.1 手法

1回の対話には3人の被験者が参加する。その際、対話に集中させる/集中させない役割に分けるため、すべての被験者には対話と同時進行で、他の対話者に気づかれないように別のタスクを行わせる。対話に集中させる被験者（Off Task 対話者）を対話者3人のうち1人、対話に集中させない被験者（On Task 対話者）を残りの2人に行わせる。Off Task 対話者には会話を通してのみ達成されるタスク（性格あてゲーム）を、On Task 対話者の2人には会話とは関係のないタスク（しりとり）を行わせる。対話の話題は予め用意した4つのお題の中から実験者が1対話システムにつき1つ選ぶ。この時、Off Task 対話者が与えた被傾聴感、Off Task 対話者から感じられた社会的スキル、Off Task 対話者の対話不安感、Off Task 対話者が与えた被傾聴感の自己評価の4点を評価する。

3.1.1 被験者

被験者として、日本語を母語とする72人の青年（男32人、女40人、20-25歳）が3人1組になって対話実験に参加した。各組に参加する3人のうち、1人はOff Task 対話者、残り2人はOn Task 対話者であった。各タスクはランダムに割り当てられた。よってOff Task 対話者は24人、On Task 対話者は48人であった。対話は2回行い、2つの対話システムを使用する順序や対話のお題はランダムに割り当てた。割り当てられたタスクは、2回の対話とも同じとした。

3.1.2 実験装置

対話は、図2に示すような、他の被験者や実験者の姿や声、周辺音が聞こえない防音室を3部屋用意して実施した。防音室内には、机と椅子が用意されている。机にはPCと卓上ライトを設置し、PCには各条件にある対話システムが表示される。実験時、被験者はオーバーイヤー型ヘッドセットを装着した。2つの対話システムを用意した。提案システムとして、2節で紹介したCommUTalkを用いた。ブラウザはChromeを使用した。ベースシステムとして、ビデオ対話システム「Zoom」を用いた。

3.1.3 実験刺激

CommUTalkに社会的装飾機能を施したもので対話する条件（CommUTalk条件）（図1）と、Zoomを用いて対話する条件（Zoom条件）（図3）を用意した。CommUTalk条件がZoom条件と異なる点は3点ある。画面に表示されているのが対話者自身の映像ではなく対話者のアバターである点、画面に表示される各対話者がそれぞれ別の場所ではなくアバターとして同じ仮想空間上で参加している点、非言語コミュニケーションを対話者自身が行う必要がなく対話者のアバターが代わりに行う点である。被験者には互いの匿名性を保ったまま対話に参加させるため、各被験者に3色（赤、青、黄）の色カードを左胸の上につけさせ、それぞれ「赤色」、「青色」、「黄色」という名前で対話に参加させた。



図2 対話システム使用時の環境



図3 Zoom条件で使用した対話システムの画面キャプチャ

それに合わせて各被験者には、CommUTalkを使った対話システムにおいて、自身のアバターの色、名前も各々同じ色のものを使用した。Off Task 対話者には対話中別タスクである性格あてゲームを行わせる。これは対話相手である他2人の性格を当てるクイズに対話終了後答えてもらうため、2人に色々質問をするというタスクである。On Task 対話者には対話中2人で常にしりとりを、スマホのチャットアプリを使用して行わせる。但しOn Task 対話者同士でしりとりを行っていることは明示していない。被験者は各条件で8分間、4つのお題（仕事/バイトと私、旅行と私、お酒と私、コロナと私）の中から実験者がランダムで1つ指定したもののについて対話する。

3.1.4 実験の流れ

3人の被験者の前に実験内容が書かれた紙を提示しながら実験者が被験者全員に対して内容を音読した。まず、実験者は被験者に、最近コロナでZoomといった遠隔でのチャットが流行っているが、そういった遠隔対話をしていると様々な邪魔が入るときがあるのでこれをいかにばれず、素早く処理できるかを調査する、と教示する。さらに、2種類の遠隔対話システムを各1回ずつ使用し対話すること、3人のうち誰か1人にタスクを与えるのであればないように対話に参加すること、タスクは重要なので真剣に取り組むことを教示した。被験者自身に、指示した色カードを左胸の上につけさせ、それぞれ「赤色さん」、「青色さん」、「黄色さん」という名前で対話に参加してもらうことを教示した。その後、被験者自身についての事前のアンケートに答えさせた。被験者全員がアンケートの回答を終えた後、実験の対話で使用する2つの対話システムのうち1回目を使用するシステムについての説明を行った。それぞれのシステムについて、CommUTalk条件では、アバターチャットシステムとは人の話し声に合わせて動くロボットを見ながら、話せるシステムであり、画面内での一番手前のロボットに乗り移って奥2体のアバターと対話を行

う、Zoom 条件ではビデオチャットシステム Zoom を用いたシステムであり、相手の顔を見ながら対話を行う、と提示した。1 回目使用する対話システムの説明後、対話を行う防音室での被験者の行動を指示した。内容は、説明したシステムで 8 分間対話を行うため対話を行う小部屋に移動してもらうこと、対話システムはすでに作動しているので、ヘッドセットをつけること、その際各被験者が対話を行う小部屋内の机の上の用紙の指示に従うこと、用紙には対話のお題と、対話を始める人、さらに被験者のうち 1 人には秘密のタスクが追加で提示していること、対話内容としては 4 つのお題の中から 1 つ提示していること、を教示した。実験者は、各被験者を対話を行う部屋に案内した。被験者は部屋内の用紙に提示された指示に従い、3 人で対話を行った。各被験者が従う用紙に記載の指示内容はそれぞれ異なる。On Task 対話者の 2 人として赤色、青色カードを身に着けた被験者にはそれぞれ以下のように指示した。

- 赤/青色さん：机の上にあるスマホで別室の方とチャットでしりとりを行ってもらいます。できる限り早く返すようにしてください。しりとりできた数と、返答の速さを測ってランキングをつけているのでできる限り早く行ってください。対話が始まったらあなたからしりとりを始めてください/相手側から始めるのでそれに続いて行ってください。今回の対話のお題は〔対話のお題〕です。黄色さんから声がかかるので、それまで待ってください。

Off Task 対話者として黄色のカードを身に着けた被験者には以下のように指示する

- 黄色さん：相手 2 人の性格を当ててもらいます。会話後に、先ほど回答したアンケート項目の中からランダムで選んだ 5 項目を提示するので、各項目についてそれぞれの相手がどう答えたか当ててください。どのアンケート項目が提示されるかはランダムなので、色々 2 人に質問してみてください。今回の対話のお題は〔対話のお題〕です。あなたから対話を始めてください。例：「こんにちは、対話を始めましょう」

なお、上記表記内容の〔対話のお題〕には全員共通して先述の 4 つのお題からランダムで 1 つ選択したものを提示した。対話の開始は、対話開始の指示を受けた、Off Task 対話者のタイミングで開始した。8 分間の対話後、実験者は被験者を対話する部屋から退出させ、実験者が教示した場所へ移動させた。その後アンケートを回答させた。Off Task 対話者には相手 2 人の性格を当てるクイズを追加で答えさせた。アンケート回答中、実験者は被験者らが 2 回目使用する対話システムの準備を行った。被験者全員がアンケートの回答を終えた後、実験の対話で使用する 2 つの対話システムのうち 2 回目使用するシステムについての説明を行った。その後、1 回目の対話システムの教示と同様、対話を行う部屋と部屋内での被験者の行動指示について教示し、被験者を対話を行う部屋に再度案内し、被験者は用紙に提示された指示に従い、3 人で対話を行わせた。8 分間の対話終了後、実験者が教示した場所へ、集合させ、アンケートを回答させた。

3.1.5 評価方法

簡易型関係認知目録（話し手用）[4] のうちの人間尊重の態度尺度を用いた。この尺度は聴き手の態度が話

し手にとってどのようなものであったかを評価するというものである。本研究の目的は対話システム使用時の On Task 対話者が与える被傾聴感を測ることである。被験者には、話し手として感じた他の 2 人の被験者の聴き手としての印象を個別に回答させた。KiSS-18 は、身についている社会的スキルの自己評価の程度を測るもの [5] であり、対話システムなどの人工物にたいしても、この尺度のスコアが影響することが知られている [6]。またこの尺度は、自己評価と他者評価に相関があることも分かっている [7]。よって本研究では、KiSS-18 を対話システムで対話した際の対話者の社会的スキルを他者評価で測るものとする。被験者には、対話システムで対話した際に感じた他の 2 人の被験者の社会的スキルを個別に回答させた。PRCA-24 は様々な対話場面における実際の、あるいは想像上の対人対話に関連した恐怖あるいは不安のレベルを測るものである [8]。PRCA-24 には集団討論、会議やミーティング、1 対 1 対話、人前でのスピーチという場面ごとに分類して各 6 項目、全 24 項目の尺度で対話不安のレベルを測っている。本研究の目的は複数人での対話における対話システム使用時の対話者の対話不安感を測ることである。よって本研究では、PRCA-24 尺度のうち会議やミーティングの場面についての尺度を用いることにした。簡易型関係認知目録（聴き手用）のうちの人間尊重の態度尺度は聴き手の態度が話し手にとってどのようなものであったかを聴き手が自己評価し、振り返るものである [4]。この尺度は、対話相手の聴き手として自身がしっかりと振舞っているのかを自己評価するものであるため、本研究では対話不安におけるもう一つの尺度として測ることができると考えた。対話とは関係ないタスクを行った量によって対話の負荷を評価することで、対話者の対話システムの対話パフォーマンスを調査した。対話中に On Task 対話者同士でやっていることは教示せずしりとりを行ってもらった。各対話システムにおける対話の負荷 8 分間の対話で出たしりとり量を両条件において計測した。

3.2 結果

CommUTalk 条件と Zoom 条件の結果を 3 つの観点から比較した。それぞれの結果の詳細を以下に示す。On Task 対話者が与えた、Off Task 対話者の被傾聴感に関し、分析手法として対応有の t 検定を行った ($n=24$)。有意水準は 5% としてある。分析の結果、CommUTalk 条件 ($M=82.7$, $SD=10.6$) と Zoom 条件 ($M=80.4$, $SD=9.54$) に有意な差は認められなかった ($t=-1.38$, $p=0.180$)。Off Task 被験者が評価した、On Task 対話者から感じられた社会的スキルに関し、分析手法として対応ありの t 検定を行った ($n=24$)。有意水準は 5% としてある。分析の結果、CommUTalk 条件 ($M=68.1$, $SD=7.46$) と Zoom 条件 ($M=67.2$, $SD=7.04$) に有意な差は認められなかった ($t=-0.971$, $p=0.341$)。On Task 対話者が評価した、対話不安に関する結果を図 4 に示す。分析手法として対応ありの t 検定を行った ($n=48$)。有意水準は 5% としてある。分析の結果、CommUTalk 条件 ($M=15.59$, $SD=4.69$) のほうが Zoom 条件 ($M=17.04$, $SD=5.07$) と比べ有意に改善していることが認められた ($t=-2.43$, $p=0.047$)。On Task 対話者が与えた、Off Task 対話者の被傾聴感の自己評価の観点からみた対話不安に関し、分析手法として対応ありの t 検定を行った ($n=24$)。有意水準は 5% としてある。分析の結果、CommUTalk 条件 ($M=79.9$, $SD=8.60$) と Zoom 条件 ($M=79.0$, $SD=10.9$) に有意

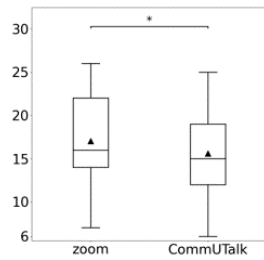


図4 Off Task 対話者の対話不安

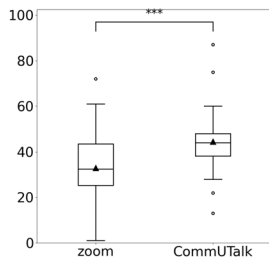


図5 On Task 対話者が行ったタスク量

な差は認められなかった ($t=-0.604$, $p=0.552$)。On Task 対話者が対話中に行ったタスク量に関する結果を図5に示す。分析手法として、対応有の t 検定を行った ($n=48$) 有意水準は5%としてある。分析の結果、CommUTalk 条件 ($M=44.6$, $SD=14.8$) のほうが、Zoom 条件 ($M=33.1$, $SD=15.7$) と比べ有意に高いことが認められた ($t=-4.20$, $p=0.0003$)。

4. 考察

結果より、On Task 対話者が評価した対話不安に関して、CommUTalk 条件のほうが Zoom 条件に比べて改善した。また、On Task 対話者が対話中に行ったタスク量に関しても、CommUTalk 条件のほうが Zoom 条件に比べ有意に増えた。被傾聴感、社会的スキルに関しては2条件間で有意な差は見られなかった。CommUTalk 条件と Zoom 条件において、対話に集中していない人が感じる対話不安において CommUTalk 条件のほうが有意に改善することが認められた。これは自身のアバターが代わりに対話相手への非言語コミュニケーションを行い、自身はそれを行う必要がなくなるため、対話者がビデオ対話システムでの対話時に行っていた、発話者のほうに目線を合わせる、相槌の行動としてうなづく、発話者として聴き手が関心を持つようにジェスチャーをするなどの非言語応答の負荷が減り、対話不安が減少したと考えられる。実際に、対話中の on Task 対話者が対話とは別に行ったタスク量は CommUTalk 条件のほうが有意に増えることが認められた理由として、上記の非言語応答の作業負荷の減少が対話のパフォーマンスが向上したためだと考えられる。一方で、対話に集中していない人が与える被傾聴感、社会的スキルに関しては2条件間で有意な差は見られなかった。その理由として対話者は、Zoom 条件の対話相手の顔や表情や、CommUTalk 条件の社会的装飾機能などの見かけだけではない部分で被傾聴感や社会的スキルを評価していると考えられる。また、社会的装飾機能は2節に示す3つの機能が含まれるが、それぞれが対話者

にもたらす効果は検証していない、どの機能がどの程度結果に貢献しているかは不明なので、検証する必要がある。

5. 結言

本論文では、遠隔対話システムとして、社会的装飾機能を持つ複数アバター対話システムを提案した。また、提案システムとビデオ対話システムの2つの対話システムを用いて被験者に3人1組で対話させ、対話に集中しづらい人が与えた被傾聴感、社会的スキル、対話に集中しづらい人の対話不安について、心理実験により評価を行った。その結果、提案システムのほうが、ビデオ対話システムに比べて、対話に集中しづらい人の対話不安感は有意に減少する、対話中の on Task 対話者が対話とは別に行ったタスク量は有意に増えることが認められた。一方で対話に集中しづらい人が与える被傾聴感、社会的スキルに関しては2条件間で有意な差は見られなかった。以上の結果から対話に集中しづらい人のタスク量の改善により、社会的装飾機能を持つアバターが代わりに非言語コミュニケーションを行ってくれ、対話により集中できることにより、ビデオ対話システムでの対話時の非言語応答の負荷が減少し、対話不安が減少したと考えられる。また、社会的装飾機能に含まれる複数の機能のうちどれがより対話者にもたらす効果が高いのかは検証していないので検証する必要があること、より豊かな社会的装飾機能、例えば複数アバター対話システムのアバターが同じ仮想空間内にいることを利用して、発話者や発話内容によってアバターのふるまいだけでなく位置などを半自律で動かすことで対話者全員の発話を促す、といった開発を進めることが今後の研究課題となる。本研究は、複数人での遠隔対話の新しい遠隔対話システムの可能性を示すものであると考える。

参考文献

- [1] 谷田貝雅典, 永岡慶三, 坂井滋和, 安田孝美. 視線一致型及び従来型 TV 会議システムを利用した2大学間遠隔交流学习. 情報文化学会誌= Journal of the Japan Information-culture Society, Vol. 18, No. 1, pp. 34-41, 2011.
- [2] 森川治, 山下樹里, 福井幸男, 佐藤滋. 超鏡対話における対話者の配置と話しやすさの関係. インタラクシオン, Vol. 2001, pp. 179-186, 2001.
- [3] 山田涼介, 由井蘭隆也, Others. 遠隔会議を活性化するためのアバター表情演出に関する提案. 研究報告デジタルコンテンツクリエーション (DCC), Vol. 2020, No. 36, pp. 1-6, 2020.
- [4] 三島徳雄. 積極的傾聴とは; 一般的な傾聴法. 三島徳雄, 新小田春美編著 『看』, 1999.
- [5] 菊池章夫, Others. KiSS-18 研究ノート. 岩手県立大学社会福祉学部紀要, Vol. 6, No. 2, pp. 41-51, 2004.
- [6] Tsunehiro Arimoto, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Multiple-robot conversational patterns for concealing incoherent responses. *Adv. Robot.*, Vol. 10, No. 5, pp. 583-593, November 2018.
- [7] 内藤諄人. 社会的スキルの自己評価と他者評価の一致について. 2013.
- [8] James C McCroskey, Michael J Beatty, Patricia Kearney, and Timothy G Plax. The content validity of the PRCA - 24 as a measure of communication apprehension across communication contexts. *Commun. Q.*, Vol. 33, No. 3, pp. 165-173, June 1985.