

複数ロボットとの位置関係がユーザの対話負荷に与える影響

松元 崇裕^{1,a)} 後藤 充裕^{2,b)} 石井 亮^{3,c)} 渡部 智樹^{2,d)} 山田 智広^{2,e)} 今井 倫太^{2,f)}

受付日 2018年4月18日, 採録日 2018年7月10日

概要: ヒューマンロボットインタラクション分野において, 人との対話に複数ロボットを用いると多くの利点があることが知られている. しかし本論文では, 我々の高齢者とロボットの対話トライアルにおいて, 従来研究に反し, 高齢者が2台ロボットとの対話よりも1台ロボットとの対話に高い評価を与えた事例を紹介する. 我々は, トライアル結果より, 人と2台ロボットの不適切な位置関係が, 高齢者の会話負荷が増やしたことが原因ではないかという仮説を立てた. 高齢者にとって, 負荷がシステムの評価に与える影響は大きく, ロボットを高齢者に利用してもらうためには負荷を減らすことが重要であることが知られている. そこで本研究は, 人と複数台ロボットの会話時の位置関係が負荷に与える影響について着目し, 仮説に基づく検証を行った. 本研究では, まず被験者が2台のロボットと, 2つの位置関係で対話を行う比較実験を実施し, NASA-TLX と二重課題法を用いて対話時の負荷評価を行った. さらに同様の位置条件で, 少人数の高齢者を実験参加者としたケーススタディを実施した. 本研究では比較実験とケーススタディを通じて, 人と複数台ロボットのグループ対話において, 人が対話相手のロボットを同時に見ることができない位置関係では, 対話相手のロボットを同時に見ることができる位置関係と比べて, 人の対話負荷が増加し, 対話に対する評価が減少することを示す. さらに位置関係における対話時の負荷は, 身体的な負荷だけではなく, むしろ心理的な負荷が主な原因となることを示す.

キーワード: ヒューマンロボットインタラクション, 対話負荷, 複数ロボット対話, 会話陣形

Spatial Arrangement Effects to Participant Workload in Human Multi-robots Conversation

TAKAHIRO MATSUMOTO^{1,a)} MITSUHIRO GOTO^{2,b)} RYO ISHII^{3,c)} TOMOKI WATANABE^{2,d)}
TOMOHIRO YAMADA^{2,e)} MICHITA IMAI^{2,f)}

Received: April 18, 2018, Accepted: July 10, 2018

Abstract: Several benefits obtained using multiple robots in conversation have been reported in the human-robot interaction field. This paper first presents pre-trial results by which elderly people assigned a lower rating to a conversation with two robots than to one with a single robot. Observations of the trial suggest the hypothesis that an inappropriate spatial arrangement between robots and humans increases the workload in a conversation. Reducing the workload is important, especially when robots are used by elderly people. Therefore, we specifically examine the workload that is influenced by the spatial arrangement in group conversation. To verify the hypothesis, we use a NASA-TLX and a dual-task method to evaluate the workload and to conduct a comparative experiment in which the participant talks with two robots in two spatial arrangements. We also conduct a case study for elderly people in the same conversational conditions. From these experiments, we demonstrate that the spatial arrangement in which people cannot see both robots simultaneously increases their conversational workload and decreases their evaluation of the dialogue compared to a spatial arrangement by which people can see both robots simultaneously. We also show that the primary cause of the workload by positioning is not physical but mental.

Keywords: human-robot interaction, conversational workload, multi-robot conversation, conversational formation

¹ 慶應義塾大学

Keio University, Yokohama, Kanagawa 223-8522, Japan

² NTT サービスエボリューション研究所

NTT Service Evolution Laboratories, Yokosuka, Kanagawa 239-0847, Japan

³ NTT メディアインテリジェンス研究所

NTT Media Intelligence Laboratories, Yokosuka, Kanagawa 239-0847, Japan

a) matumoto@ailab.ics.keio.ac.jp

b) gotou.mitsuhiro@lab.ntt.co.jp

c) ishii.ryo@lab.ntt.co.jp

d) watanabe.tomoki@lab.ntt.co.jp

e) tomohiro.yamada.mr@hco.ntt.co.jp

f) michita@ailab.ics.keio.ac.jp

1. 序論

人との会話を通じてサービスするロボットにおいて、複数のロボットを対話へ使う利点が着目されている。たとえば、人が2台のロボットと対話するときには、1台のロボットと対話するときと比べて話の一貫性や親密さをより感じることが報告されている [1]。またロボットどうしの会話をあらかじめ見た人は、そのロボットの発話をより理解しやすくなり、ロボットとの対話も自然かつスムーズに感じることが分かっている [2]。複数台ロボットを用いた対話システムに関する研究は、1台ロボットを対象としたものと比べて数は少ない。そのため、上記の利点を生かしたサービスを実現するためにも、複数ロボットの対話システムにおける知見を蓄積していくことは重要である。

一方で、対話においては、位置関係がコミュニケーションへ大きな影響を与えることが知られている。Hall は、提唱する近接学の理論において、コミュニケーションにおける距離は密接距離、個体距離、社会距離、公衆距離の4段階が存在し、各段階で人の振舞いや言葉遣いに変化することを示した [3]。また Kendon は対話をする人々の間に一定の空間が保たれることを発見し、対話参加者がその空間を維持する仕組みを F-陣形として定義した [4]。F-陣形には複数の形状が存在し、F-陣形の形状と参加者の位置により、対話への参加態度が変化することを示した [5]。

ヒューマンロボットインタラクション (HRI) の分野でも、近接学や F-陣形の考えに基づき、ロボットと人の位置関係に関する様々な調査が行われてきた [6], [7], [8], [9]。上記の従来研究により、ロボットの位置は、人と同じようにコミュニケーションへ影響を与えることが分かっている。通常ロボットは実体を持つため、人とロボットの間には何らかの位置関係が生じる。すなわち対話ロボットの開発者やデザイナーが意図を持って、人とロボットの位置関係をデザインしたか否かにかかわらず、ロボットは人へ位置関係に基づく何らかの影響を与えてしまう。そのため位置関係が人に与える影響を解き明かし、その知見を利用できるようにすることは、人とロボットがより望ましいコミュニケーションを実現するために重要である。

望ましいコミュニケーションの1つの要素に、情報伝達で相手に疲れを感じさせないことがあげられる。特に高齢者にとっては、簡単に利用可能であることが、そのシステムを受容するか否かに大きく影響 [10]、若年層と比べて身体的・心理的負荷を減らすことがより重要であることが分かっている [11]。したがって対話ロボットが高齢者に利用してもらうためには、対話時の負荷を減らし、相手を疲れさせないことが重要である。ロボットとの対話の負荷を減らす研究には、対話戦略に着目した研究 [12], [13] や、対話時のジェスチャ [14] に着目した研究などが実施されている。しかしながら、対話時の位置関係が対話負荷へ与える

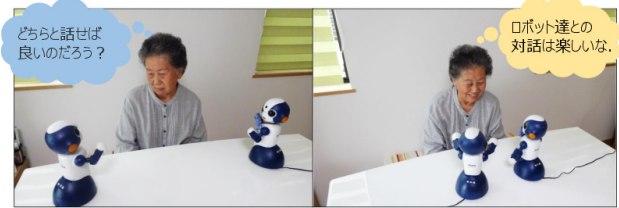


図 1 複数台ロボットとの対話において人に高い対話負荷 (左) と低い対話負荷 (右) を与える位置関係の例

Fig. 1 Spatial arrangements that support high workload (left) and low workload (right) in conversation.

影響については、現在までほとんど議論されてこなかった。

本論文では、人と複数台のロボットの会話において、位置関係の違いが対話負荷へ与える影響について着目する。最初に本論文では、高齢者が1台のロボットと、2台のロボットの、2条件で対話をしたトライアルにおいて、高齢者が2台のロボットより、1台のロボットとの対話を高く評価した結果を紹介する。次に、トライアルの結果から、両方のロボットを同時に見ることができない位置関係にいる人は、両方のロボットを同時に見ることができる位置関係にいる人と比べて、高い対話負荷を感じるという仮説を示す。図 1 は、仮説における人と2台ロボットの位置関係を示した例である。我々は仮説検証のため比較実験とケーススタディの2つの実験を行った。本論文では2つの実験の詳細および結果と、対話時の位置関係が対話負荷と対話への印象に与える影響について考察を述べる。

本論文の、人とロボットのインタラクション理解の分野への貢献は次の3点である。第1に、本研究は、人と複数台のロボットの位置関係において、人が両方のロボットを同時に見ることができる位置関係をとった場合には、そうでない場合と比べ会話時の負荷が増加することを示す。第2に、高い対話負荷を人に与える不適切な位置関係は、その位置関係に置かれた人の対話への評価を減少させてしまうことを示す。第3に、不適切な位置関係がもたらす対話負荷は、顔を動かすことによる身体的な負荷だけでなく心理的負荷が影響しており、さらに身体的負荷よりも心理的負荷が主要因であることを示す。本研究は、人と複数台ロボットの対話において位置関係が、人の対話負荷へ与える影響を調査した初めての研究である。対話ロボットの開発者やデザイナーが、対話に及ぼす影響を考えて位置関係を適切にデザインするために、本研究成果は重要な1つの知見になると考える。

2. 関連研究

2.1 複数ロボットとの対話研究

HRI 分野において、人がロボットと対話するとき、1台よりも複数のロボットと対話する方が、対話やロボットの印象へ良い影響を与えることが示されている。たとえば、

Kanda らはロボットどうしの会話を人に事前に見せると、人はそのロボットと対話するときに、ロボットの発話理解が容易になり、ロボットとの対話の自然さや円滑さが向上することを示している [2]. Hayashi らは、社会的受動メディアというロボットどうしの会話を通して人へ情報提示するコンセプトを提案し [15], 2 台のロボットによる情報提示が 1 台による情報提示よりも有意に注目を引くことを示した [15]. Marynel らは、パートナーロボットだけが理解できる独自言語を話すランプ型ロボットであっても、ランプ型ロボットの存在が、パートナーロボットの発話に人の注目をより集めることに貢献し、ロボットとの対話の楽しさ向上に寄与することを示した [16]. Iio らは、2 台ロボット用の対話戦略を提案し、人が 1 台のロボットとの対話より、2 台のロボットとの対話の方が、ロボットの発話に一貫性を感じ、ロボットと親密さを強く感じることを示した [1].

2.2 会話時の位置関係を扱った関連研究

Kendon はビデオカメラで撮影した対話の様子から、対話時に各個人が自身の手前に広がる操作領域 (transactional segment) を維持するように空間配置行動をとることを発見し、F-陣形という概念を提唱した [4]. F 陣形は 2 人以上が集まって会話するときに、O 形の空間 (O-Space) を取り囲むように、対話者が下半身を向けて身体配置することで構成される. 2 人で構成される F-陣形には、Fig.1 (上) で示した vis-à-vis, L-shape, side-by-side などの陣形がとられることが知られている.

F-陣形は人とロボットの間に形成され、対話時の状況によって構成される F-陣形の頻度が変化することが知られている [17]. そのため HRI 分野においても、F 陣形の概念に基づき、対話時の人とロボットの空間配置に関する研究が行われてきた. たとえば Yamaoka らは、環境中の対象を説明するときに、ロボットが F-陣形の概念に基づいて位置と体の向きを決定することが有効であることを示した [8], [9]. また Kuzuoka らは人がロボットに対しても F-陣形を維持するように行動することを利用して、ロボットの下半身のねじりにより、人の位置を誘導できることを示した [22]. Chao らは対話開始から F-陣形を構築するまでに着目し、ロボットが人と自然にインタラクションをスタートする方法を提案した [38]. 上記の従来研究では、主に人とロボットの 1 対 1 のインタラクションにおける位置関係について着目している.

一方、F-陣形は 3 人以上の場合にも構成され、2 人で構成される F-陣形と比べて、より複雑となることが知られている [4]. F-陣形が 2 人で構築される場合、対話参加者同士はつねに対象な位置関係となるが (図 2 : 左上), 3 人以上の場合には Circle のように全参加者の方向関係が対象な陣形 (図 2 : 左下) 以外にも、Semicircle や Rectangle といった非対称な陣形 (図 2 : 右下) が存在する. 非対称な陣形で

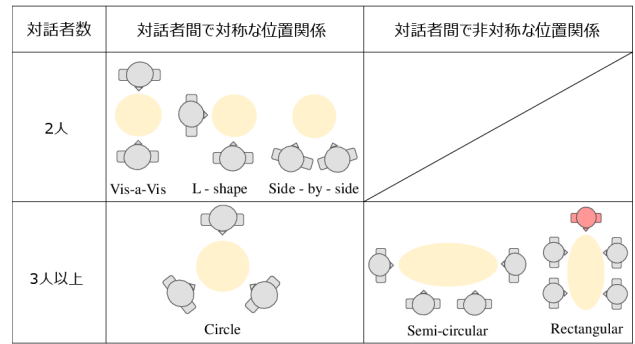


図 2 対話参加者の人数と位置関係による F 陣形の分類

Fig. 2 Classification of F-formation configurations by participant number and spatial relation.

生じる興味深い効果に、参対話における発話権 (Speaking rights) への影響があげられる [5]. たとえば陣形が Circle である場合、各参加者の発話権は平等であり参加者の発話量は偏らない傾向にあるが、陣形が Rectangle の場合には図 2 の赤色で示された人物が強い発話権を持ち、他参加者より長い時間発話する傾向のあることが分かっている.

3 人以上で構築される F-陣形を対象とした研究には、カメラ映像から F-陣形を検出する研究 [18], [19] や、すでに構築された F-陣形へロボットがどの位置から加わるべきかということに文化的差異があることを示した研究 [20] などが行われてきた. しかし非対称な陣形における参加者の位置が、対話の印象や、参加者自体にどのような影響を与えるかについては、ほとんど調査されていない.

2.3 対話時の負荷についての関連研究

対話システムの負荷についての研究は、運転や高齢者との対話という領域で、複数の研究が実施されてきた. 運転中の対話負荷についての研究では、対話内容や話しかけるタイミングによって、対話中の運転者の心理的負荷の量が変化することが分かっている. Christopher らは、運転中の電話において、対話内容が、電話の利用方法 (ハンズフリー vs. 手で持つ) よりも、運転タスクに大きな影響を与えることを示した [21]. また Kennington らは、音声によるドライバーへの情報提示システムにおいて、運転データを元に発話提示タイミングを適切に制御すると、ドライバーの認知負荷を下げられることを示した [22].

また従来研究では対話ロボットを用いたドライバーへの情報提示における負荷に着目した研究も行われてきた. たとえばダッシュボードに設置された小型ロボットとの対話は、電話よりもドライバーに与える心理的負荷が小さく、同乗者との会話と同程度の負荷であることが示されている [23]. また Okada らは複数のロボットどうしの会話を介して人へ情報提示すると、人にロボットへの返答を強制することを避けられるため、認知負荷が軽減されることを示し [24], さらに複数ロボットによるドライバー向けの負荷の小さい情

報提示手法を提案している [25].

高齢者にとっても、会話システムを利用時の負荷を下げることは重要である。高齢者は老化のため、情報処理速度、心理的な柔軟性、流動的知性、記憶など様々な認知機能が低下することが知られている [26]. 従来研究でも、高齢者向けの情報システムは、心理的な負荷を下げることに特に重要であることが言及されており [11], [27], 高齢者は、若者と比べて、ユーザビリティが、システム全体の評価へ大きく影響することが示されている [28].

高齢者とロボットの対話において、心理的負荷を下げる方法として、従来研究では主に対話戦略からのアプローチが行われてきた。たとえば、ロボットが回答の選択肢を狭めるように質問し、会話内容をつど確認する対話戦略 [13] や、対話で利用する単語の一貫性を保つ対話戦略 [12] が提案されている。また対話戦略以外では、ロボットが発話にジェスチャを交えて情報を伝えると、理解の難しい内容に限り音声のみで伝えるよりも認知負荷を減らせることが Manja らによって示されている [14].

対話時の負荷についての多くの従来研究では、発話内容やタイミングに焦点を当てており、ロボットの身体性と対話時の負荷、特にロボットとの位置関係と認知負荷の関係についてはほとんど明らかになっていない。本研究では複数ロボットと人の位置関係が対話時の負荷に与える影響に着目し、次章のトライアルでの事例を通じて人とロボットの位置関係によっては、対話に複数台のロボットを用いてもその利点が発揮されない可能性について示す。さらに2つの実験を通じて、対話時の人と複数ロボットの位置関係が、人の対話負荷へ影響を与えることを明らかにする。

3. 高齢者向けリハビリテーション施設におけるトライアル

我々はロボットと高齢者の対話トライアルを実施した。トライアルは会話負荷と対話位置の関係調査を直接目的としていない。しかし本トライアルの結果は、対話時にロボットが不適切な位置関係をとることが、人の対話負荷へ与える影響について重要な示唆を含んでいる。そこで本章ではトライアル内容と結果を紹介し、対話時の位置関係と負荷についての仮説を示す。

トライアルは認知症予防のためのロボット対話システムを評価する目的で実施された。実験はロボット2台と高齢者1人の対話と、ロボット1台と高齢者1人での対話の2条件で実施され、どちらが高齢者にとって望ましいか比較評価が行われた。トライアルへは普段身体機能のリハビリで施設を利用中している高齢者14人（男性7人、女性7人）が参加し、日常会話には問題がなく、ロボットとの会話が未経験である人が選ばれた。参加者の平均年齢は76.5歳（SD = 9.58, min = 59, max = 91）であった。実験参加者のうち4名は、厚生労働省における指標において、買

物・家事・服薬管理などの IADL (Instrumental Activities of Daily Living) で一部支援が必要とされる、要支援1または2に認定されていた（要支援1:1名、要支援2:3名）。残りの10名の参加者は、IADLに加えて食事・排せつ・入浴などの ADL (Activities of Daily Living) においても一部介護が必要な、要介護1・2および3に認定されていた（要介護1:3名、要介護2:6名、要介護3:1名）。

実験には Viston 社製の小型ロボット Sota (280 mm (H) × 140 mm (W) × 160 mm (D)) を用いた。ロボットの自由度 (DoF) は首が3DOF、肩が1DOF、肘が1DOF、胴体が1DOFの計6DOFで、顔方向や手足を動かすことによる簡単なジェスチャ表現が可能である。ロボットの視線方向は、顔方向により表現の代替可能という従来研究 [29] から顔方向で表現した。また Sota は口に該当する箇所に赤色の LED ライトがついており、音声再生中は発話中の表現として口元の LED を点滅させた。

ロボットと高齢者の対話は、あらかじめ人手で作成された対話シナリオに沿って実施された。対話シナリオには、高齢者の若齢期の記憶を喚起する目的で「昔の遊び」「小学校の給食」「昔の家事」「昔のテレビ番組」の4テーマを用いた。対話は前後半の2つのセクションで実施され、「昔の遊び」「小学校の給食」が前半、「昔の家事」「昔のテレビ番組」が後半の対話テーマとした。

両条件の比較のため、対話シナリオはロボット1台用のシナリオと、ロボット2台用のシナリオが用意された。両条件での対話シナリオの主な違いは、同じ対話内容を1台のロボットに話させるか、2台のロボットが分担して話すかという違いであり、対話内容は可能な限り等価となるよう作成された。トライアルは参加者内実験で実施され、前後半をロボット1台条件と2台条件のどちらで行うかについては、参加者ごとにカウンターバランスを取った。対話中は高齢者の記憶喚起を促す目的で、テーマごとに関連する写真が1枚、参加者の前のモニタに表示された。対話シナリオ設計の詳細、およびロボットの視線制御やジェスチャ制御は4章で詳細を述べる。また各条件における対話時間は約10分であった。

対話後には、各条件を評価する7段階のリッカート尺度によるアンケート（低評価:1, 高評価:7）を用いて、各条件における対話の印象評価を行った。アンケートには4つの単項目、「Q1:どのくらい(2台の)ロボットと会話が弾んだか」「Q2:どのくらい(2台の)ロボットとの対話を楽しみ感じたか」「Q3:どのくらい(2台の)ロボットと会話するのが簡単だったか」「Q4:どのくらい(2台の)ロボットと次回もまた話をしたいと思うか」を用いた。

複数ロボットを用いた対話システムのメリットに関する従来研究 [1], [2], [15], [16], [30] より、我々は1台ロボット条件よりも2台ロボット条件の方が高い評価になると予想していた。しかし図4に示すように、アンケート結

果は全項目で1台ロボット条件が2台ロボット条件より有意に高い評価となった。検定にはウィルコクソンの符号つき順位検定を用い、結果はQ1 ($z = 2.36, p = 0.018$), Q2 ($z = 2.38, p = 0.017$), Q3 ($z = 3.33, p < 0.01$), Q4 ($z = 3.32, p < 0.01$)である。実験後の参加者へのインタビューでも、アンケート結果を裏付けるように「2台のロボットの相手は疲れてしまった」、「ロボットが2台いるときにはどちらと話をするのか分からなかった」など、2台ロボットへ否定的なコメントが多くあげられた。

2台ロボット条件が低評価となった原因を探るため、我々は対話中の被験者の様子についてビデオ観察を行い、条件間で参加者の視線の動きが異なることに着目した。2台ロボット条件では図3(左)に示すように、ロボットは写真提示用モニタの両隣に配置された。そのためロボットの位置は参加者の正面から左右に大きく開いており、2台のロボットを同時に視野に入れることができなかった。その結果、参加者はロボットの話の聞くとときや、ロボットに話しかけるときに、左右のロボットを交互に見る動作が頻繁に観察された。一方、1台ロボット条件では、被験者から見てモニタの右側に置かれたロボットのみとの対話となる。そのため高齢者はモニタの写真を見るとき以外は、ロボットのみを見て会話しており、顔動かして視界を切り替える動作はほとんど観察されなかった。

上記の観察より、我々はグループ対話時の陣形における位置関係と参加者の認知負荷について次の仮説を立てた。

仮説: 複数参加者の対話において、他参加者を視認するために顔方向の切り替えが必要な人は、顔方向を変えずと

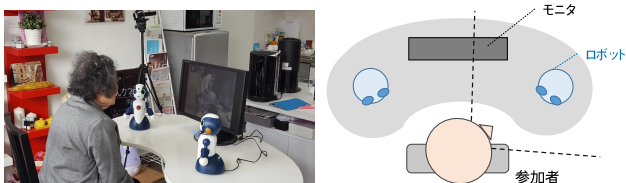


図3 トライアルでの2台ロボット条件の対話の様子(左)と位置関係の俯瞰図(右)

Fig. 3 Pre-trial of two robots condition (left) and a top view to show the orientational relation (right).

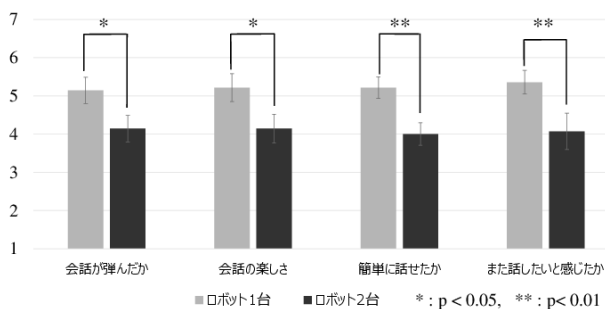


図4 プレトライアルにおけるアンケート結果の平均値

Fig. 4 Mean scores of the questionnaires administered pre-trial.

も他参加者の様子を同時に視認可能な人と比べ、対話の負荷が大きくなる。

仮説検証のため、本研究では2台ロボットと人の対話を、2台が人の視界に同時に入る位置関係と、同時に入らない位置関係の2条件で行い、人の対話負荷が条件間で変化するか調査する。

4. 複数台ロボットを用いた対話システム

本章では実験で使用した複数ロボットの対話システムを紹介する。本対話システムは一部を Wizard of Oz (WoZ) 法 [16] で制御している。ジェスチャー・視線方向・発話内容およびタイミングの制御は自動的に決定されるが、人の発話終了検知と、ロボットうなずきタイミングは、人よりシステムに通知される。

4.1 対話制御

対話の内容は、人の対話負荷へ影響を与える要因の1つになる [21]。しかし本研究は対話時の位置関係が対話負荷へ与える影響を明らかにすることを目的としている。そのためロボットには事前に定められたシナリオ(対話シナリオ)に沿って同じ内容を発話させ、対話内容の参加者間の条件統制を行った。対話シナリオは、図5に例示されるように、ロボットと参加者の両方の対話番が記載されている。

対話シナリオに基づくロボットの対話制御は次のように実施された。まず対話が始まると、ロボットは対話シナリオの各発話文を1秒の間隔をあけて合成音声の再生を行う。そしてロボットの音声再生が、対話シナリオ上で参加者の対話番まで進むと、2台のロボットは参加者の発話終了まで次の音声再生を待機する。参加者の対話番では、ロボットはシステム操作者(Wizard)による「うなずき」と「参加者発話終了」の指示を受け付け、Wizardが「うなずき」の指示を送ると、両ロボットはあらかじめ設定された「うなずき」モーションを実行する。また Wizard より「参加者発話終了」の指示が送られると、ロボットは再び対話シナリオに沿って発話文の音声再生を行う。

4.2 ジェスチャー制御

ロボットには手を振る・首を横に振る・首をかしげるな

対話者	発話文
...	...
ロボット1	最近の子供達はどんな遊びをしているか知ってる?
ロボット2	えーと... テレビゲーム? 何も思いつかないや。
ロボット2	子供の頃に家の中や、外でどんな遊びをしていましたか?
参加者	NULL [参加者の対話番]
ロボット2	なるほど! 色々な遊びを知っているんですね。
ロボット1	子供達と盛り上がるために、教えてもらった遊びを勉強しよう!
...	...

図5 本システムの対話シナリオ例

Fig. 5 Conversational scenario example of this system.

ど、19種類の短いジェスチャ（平均1.84秒、最大2.7秒、最小0.9秒）があらかじめ設定されている。各発話文で、どのジェスチャが実行されるかは下記の方法で決定された。まず本システムは、対話シナリオの各発話文に対し対話行為タグの分析を行う。対話行為はDAMSL (Dialog Act Markup in Several Layers) [31]に規定される、対話意図を抽象化したラベルセットである。本研究では文献[32]の対話行為推定アルゴリズムを用いて、各発話文へ32種類の対話行為の中からラベル付けを行った。次に各対話行為ラベルに対して割り当てるジェスチャが定義された対応表を用いて各発話文へのジェスチャの付与を行った。対話行為とジェスチャの対応表は、ロボットのジェスチャ作成者により定義された。

一方、参加者の発話番では、ロボットは傾聴の表現として、うなずき動作を実施した。動作タイミングはWizardによって決定され、Wizardが指示を送ると2台のロボットは同タイミングでうなずき動作を実施した。

4.3 視線方向制御

グループ会話において、視線は発話番の調整[33]や注目する対象の表出[34]など重要な役割を果たすことが知られている。そこで我々はグループ会話の視線データから、視線遷移の確率モデルを作成し、確率モデルに従ってロボットの視線方向を制御を行った。

視線データには、我々の従来研究で収集した、4人組の12チームが其々10分間対話したときのデータを用いた[35]。視線データは3人の注釈者により、現話者(S: Speaker)への視線、次の話者(N: Next Speaker)への視線、現話者でも次の話者でもない人(L: Listener)への視線、その他の対象(X: Other)への視線の4種類でラベル付けされている。個々の話者を決めるため、1つの発話は200msecの沈黙を区切りとして文献[36]データ収集され、各発話単位で各参加者へ現話者(S)、次の話者(N)、現話者でも次の話者でもない(L)のどの状態にあるか、参加者の役割ラベルが付与された。また各発話から、次の発話への遷移において、同話者が連続する場合にはTurn Keep (TK)、話者番が変更されるTurn Change (TC)の、対話状況ラベルが付与された。発話終了時の視線の動きは、グループ対話で次の発話番決定に特に関連することが複数の従来研究で示されている[37], [38], [39]。そこで視線制御のモデル化のため、発話終了前1,000msecから終了後200msecまでの視線方向の変化を、1つの視線遷移パターン(GTP: Gaze Transition Pattern)として束ねて定義した[35]。

表1は作成された視線方向の確率モデルである。確率モデルは、各発話番において、各対話状況(TKまたはTC)にあるときに、各参加者のGTPの発生確率を、役割(S, N, L)ごとに表したものである。制御に使用するGTPはモデルをシンプルにするため、発生頻度が少ないものを取

表1 各対話状況、各役割における視線遷移パターン(GTP)の発生確率および発生時間のモデル

Table 1 Occurrence probability and timing of the gaze transition pattern (GTP) in each turn state and role.

対話状況	役割	GTP	p	開始時間		視線時間	
				Avg.	SD	Avg.	SD
TC	S	L	.280	-2.37	2.79	4.12	3.74
		L-N	.067	-2.61	3.84	L 2.28	4.35
						N 3.82	8.03
		N	.461	-3.41	4.21	6.83	6.37
		X	.192	-2.10	3.08	4.51	4.87
		L	L	.076	-1.86	3.31	3.26
	L	N	.191	-2.91	3.96	7.27	7.04
		S	.457	-4.52	7.05	5.83	7.38
		S-N	.074	-2.75	2.28	S 2.44	2.26
						N 2.43	2.66
		X	.200	-3.09	4.11	5.70	5.77
		N	L	.150	-2.12	1.92	4.71
TK	S	S	.495	-4.42	5.22	6.77	6.12
		S-L	.078	-2.67	1.80	S 2.35	1.82
						L 1.56	2.02
		S-X	.081	-3.83	5.11	S 3.57	5.15
						X 1.97	2.08
		X	.195	-2.69	3.89	4.65	5.14
	L	L	.548	-2.41	3.38	5.09	5.95
		L-L	.101	-1.94	2.16	L 1.59	2.16
						L 2.03	2.77
		L-X	.100	-1.97	4.09	L 1.67	4.04
						X 2.49	3.02
		X	.242	-2.78	3.41	5.29	6.07
L	L	.108	-1.99	2.65	3.54	3.74	
	S	.701	-6.86	7.75	13.6	12.1	
	X	.192	-2.78	3.34	5.57	5.33	

り除き、全GTPの発生確率が5パーセント以上となるようにした。表1において、開始時間は各GTPが発話終了から何秒後に開始されるかを示しており、視線時間はGTPにおける各視線方向の継続時間の平均値(Avg.)と標準偏差(SD)を示している。確率モデルにおいて対話状況がTKの場合は、次の話者が現話者と(N=S)となるため、次の話者(N)の項目は存在しない。

表1の確率モデルと対話シナリオに基づき、我々は次のようにロボットの視線方向の制御を行った。まず対話シナリオの話者情報に基づき、各発話においてロボットと参加者がS, N, Lのうち、どの役割に該当するか決定する。さらに次発話の話者情報を参照し、現発話がTCとTKの、どちらの対話状況に該当するか決定する。次にロボットが実行するGTPを、対話状況と各ロボットの役割におけるGTPの発生確率pに基づき決定する。続いてGTPに基づき各視線制御の開始タイミングを開始時間および視線時間により決定する。開始時間および視線時間の決定では表1の平均値および標準偏差を持つ正規分布を仮定した。ただ

し各視線決定には、アクチュエータの制約のためロボットの首向き制御に最低限必要となる 200 msec 以下の遷移を含まない切断正規分布を使用した。

ロボットの発話時間は発話文から一意に定まる。そのため上記の視線方向制御は、対話シナリオから静的に決定されている。ただし参加者の発話時間は一意に定まらないため、参加者の発話番終了時刻が視線制御開始の基準となる。参加者の発話番の次の視線制御は、Wizard が参加者の発話終了をロボットへ通知した直後に実施するものとした。

5. 比較実験

5.1 Spatial Formation Arrangements

仮説検証のため比較実験とケーススタディの2つの実験を実施した。両実験の対話は、実験参加者が2台のロボットを同時に視認可能な条件と同時に視認不可能な条件の2つの位置関係で行われた。図6へ2つの位置関係を示す。

視覚内条件

視覚内条件では、2台のロボットは参加者の正面方向から、それぞれ左右10度のところに設置された。人が知覚や認知タスクに利用可能である有効視野は約20度程度であるため[40]、本位置関係は参加者が顔の向きを変えずに両ロボットを視認可能な条件となる。

視覚外条件

視覚外条件では、2台のロボットは参加者の正面方向から、それぞれ左右60度のところへ設置された。本条件下では、参加者が1台のロボットを有効視野内で見ているとき、もう1台のロボットは参加者の視線方向正面から水平方向

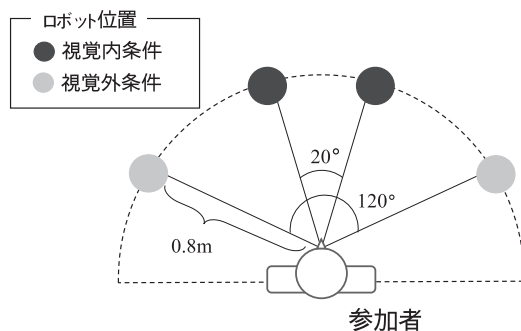


図6 実験で比較を行った位置関係

Fig. 6 Spatial arrangement in both conditions.

に100度以上の場所に位置することになる。その角度は人の周辺視野を超えるため、本位置関係は参加者は2台のロボットを同時に視野に入れることができない条件となる。

人とロボットの対話時の距離は、多くの人が0.45~1.2mの範囲を適切と感じる従来研究[17]より、本研究では両条件ともに0.8mに設定した。またロボットの身体方向は、両条件とも参加者の方向が正面となるよう設置された。参加者は水平方向のみ方向を変更可能な椅子に座り、体の向きの変更は可能だがロボットとの位置関係は変更できない条件で対話を行った。

5.2 対話負荷の測定手法

対話時の負荷を主観・客観の2つの側面で測定するため、比較実験ではNASA Talk Load Index (NASA-TLX) [41]と二重課題法[42]の2つを測定手法として用いた。

NASA-TLXは負荷の主観的評価手法として広く利用されており6つの評価尺度より構成される。表2に各評価尺度と質問内容を示す。参加者は各尺度に対して負荷の大きさを0(低い/良い)~100(高い/悪い)で評価した。

二重課題法は、負荷を評価したいタスク(メインタスク)とパフォーマンスを客観的に評価可能なタスク(サブタスク)を同時に実施させ、サブタスクのパフォーマンス結果をもって、メインタスクの負荷を評価する手法である。二重課題法はメインタスクとサブタスクを同時に実施した場合に、メインタスクの負荷の大きさがサブタスクのパフォーマンスへ影響を与えるという考えに基づいた評価である。従来研究では、加減算を解くタスク[24]、視覚信号に応じてブレーキを踏む課題[43]、しりとり課題[44]などがサブタスクに用いられた。しかし上記のサブタスクは視覚または聴覚を利用する必要があり、ロボットとの対話を直接阻害してしまう問題がある。そこで我々は触覚を用いて実施可能なサブタスクを作成した。

触覚を用いたサブタスクの実現のため、本研究では4つの振動パターンを発生させる端末を用いた。振動パターンはSingle, Double, Triple, Quadrupleから構成され、振動パターン名は20 msecの間隔で150 msecの振動が何回発生するかを示している。参加者は端末を手に持ち、SingleとTripleの振動パターンが発生したときは端末をタップ、

表2 NASA-TLXの評価尺度と質問内容

Table 2 NASA-TLX Scales and Descriptions.

評価尺度	エンドポイント	質問内容
精神的要求 (MD)	低い/高い	ロボットと話をするときに、どの程度精神的に集中する必要がありましたか?
身体的要求 (PD)	低い/高い	ロボットと話をするときに、どの程度身体的な疲れを感じましたか?
時間的圧迫感 (TD)	低い/高い	ロボットと話をするに、どの程度焦りを感じましたか?
作業達成度 (OP)	良い/悪い	どの程度、ロボットと上手く会話することができたと思いますか?
努力 (EF)	低い/高い	ロボットと話をするときに、どの程度努力することが求められましたか?
不安 (FR)	低い/高い	ロボットと話をするときに、どの程度ストレスが溜まりましたか?

Double と Quadruple の振動パターンのときは何もしないサブタスクが与えられた。サブタスク中は2秒から4秒までのランダムな間隔で、4種類の振動パターンがランダムに生成される。振動パターンの生成間隔をランダムとしたのは、参加者が振動の発生タイミングを覚えてしまうことを避けるためである。パフォーマンス評価には振動への平均応答時間および応答精度を用いた。

5.3 被験者および実験手順

実験条件の事前検証により、NASA-TLX と二重課題法による負荷測定を、トライアルを実施したりハビリ施設を利用する高齢者を対象に実施することは、質問項目やサブタスクの複雑さより困難であることが分かった。そこで比較実験には参加者に学生を用い、高齢者を対象する調査はインタビューや簡易なアンケートの評価を主とするケーススタディを比較実験を補足する目的で実施した。

比較実験には日本語を母国語とする大学生28人(男性19人・女性9人)が参加し、平均年齢は22.6歳(SD = 1.77, 19歳~27歳)であった。参加者には事前にコンピュータと会話ロボットの利用経験についてアンケートを行った。アンケートの結果、全参加者は日常的にコンピュータを利用しており、会話機能を持つロボット(おもちゃ・人形を含む)の所持者はいなかった。対話ロボットの利用経験では28名中14人が利用経験がなく、13名が過去1~2度利用したことがあり、1名が月に1~2度利用していた。またロボットとの対話についての興味を7段階のリッカート尺度(1:まったく話をしたいと思わない~7:非常に話をしたいと思う)で回答させた結果、平均値は5.28(SD = 0.81)であった。

参加者にはサブタスクについて実験前に説明を行い、タスクの慣れの影響を少なくするため参加者が問題なくタスクを遂行可能と感じるまで事前練習を実施した。説明においてはサブタスクがロボットとの対話へ影響を与えることを避けるため、対話とサブタスクを同時に実施することが困難なときには対話を優先するよう教示した。

実験は参加者内実験(視覚内条件 vs. 視覚外条件)で行い、対話ロボットには Sota を用いた。対話シナリオは前後半のパートに分けられ、各パートは同程度の長さを持ちロボットの発話時間は前半が337秒で後半が339秒であった。各位置条件での対話が前後半のどちらで実施されるかは、順序の影響を避けるためカウンターバランスをとった。またロボットの動作や会話タイミングの影響を避けるため、同パートでのロボットの視線対象・ジェスチャ・会話タイミングは条件間で同じとした。対話シナリオは比較実験とケーススタディで共通のものを用いた。下記に対話シナリオの流れを示す。

オープニング(前半): 2台のロボットが挨拶と自己紹介を行う。そしてロボットは小学校で子供達の先生をするこ

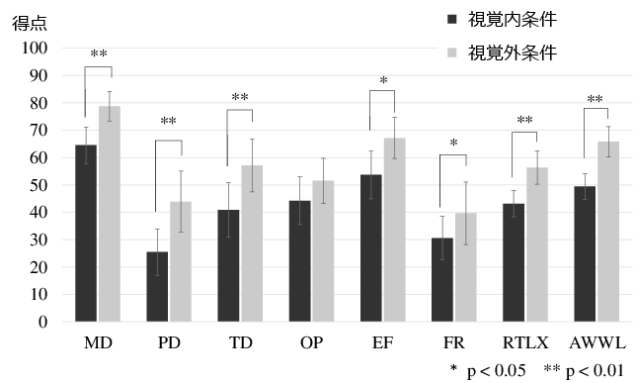


図7 両条件における NASA-TLX の平均値
 Fig. 7 Mean values of the NASA-TLX in both conditions.

とになったが、小学生のことをよく知らないため、参加者に小学校の頃の生活を教えて欲しいことが伝えられる。

メイン(前半): 小学生の様々な遊びについて会話する。会話の中で小学生のときにどのような遊びやスポーツをしていたか、ロボットは小学生とどんな遊びをするのが良いかなどの質問がロボットから参加者へ行われる。

クロージング(前半): ロボットから質問の回答への感謝と休憩を入れたいことが伝えられる。

オープニング(後半): 参加者に再び挨拶を行う。そして話の導入として、ロボットが以前は病院で患者の話し相手として働いていたことが参加者に伝えられる。さらに患者と、よく食べ物で盛り上がったことが伝えられる。

メイン(後半): 小学生の給食について会話する。会話の中で小学生はのときに給食でどのような食べ物がでたか、どのような食べ物がクラスで人気があったかなどの質問がロボットから参加者へ行われる。

クロージング(後半): ロボットから回答に対する感謝と対話の終了が伝えられる。

サブタスクは各パートのオープニングにおけるロボットの最初の発話開始から、クロージングの最後の発話終了まで実施された。また前後半の対話が終了した後、NASA-TLX とロボットの社会的受容性についてアンケートが行われた。ロボットの社会的受容性についてのアンケートは、Easy to Use, Perceived Enjoyment, Intention to Use の3つを評価する質問群から構成されるもので[45]、会話の印象評価を目的として実施された。実験後は視覚内条件と視覚外条件で感じた差について、参加者へインタビューが行われた。また被験者の前方には Kinect (Microsoft Corp.) が設置され、対話中の被験者の顔や肩の方向情報が取得された。

5.4 比較実験の結果

NASA-TLX の評価結果

図7に比較実験における NASA-TLX の各評価項目の平均値と、全評価項目の平均値(RTLX)、および Adaptive Weight Work Load(AWWL)の結果を示す。AWWLは各

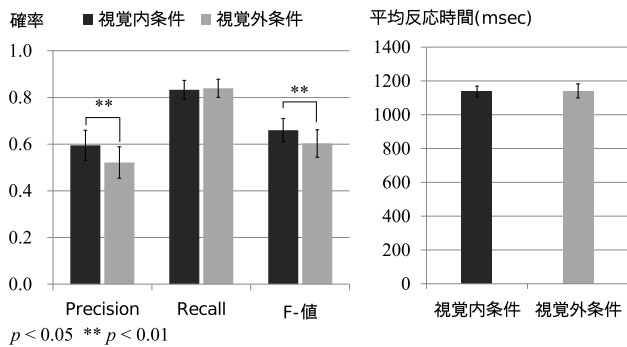


図 8 二重課題法における結果：反応の精度 (左), 平均反応時間 (右)
 Fig. 8 Dual-task performances: response accuracy (left) and mean response time (right).

評価項目のスコアの高さの順に 6 から 1 の重みをつけて平均した値であり, NASA-TLX における Weighted Workload (WWL) 得点と高い相関を示す値である [46].

精神的要求 (MD)・身体的要求 (PD)・時間的圧迫感 (TD)・努力 (EF)・不安 (FR) の 5 項目は t 検定の結果, 視覚外条件が視覚内条件と比べて, 有意に高い値となった ($MD : t(27) = 3.65, p < 0.01$), ($PD : t(27) = 4.70, p < 0.01$), ($TD : t(27) = 2.83, p < 0.01$), ($EF : t(27) = 2.64, p < 0.05$) ($FR : t(27) = 2.11, p < 0.05$). 一方で, 作業達成度 (OP) の項目は, 両条件で有意差が生じなかった ($OP : t(27) = 1.14, p = 0.26$). RTLX ($t(27) = 3.73, p < 0.01$) および AWWL ($t(27) = 4.55, p < 0.01$) は, 視覚外条件が視覚内条件と比べ, 有意に高い値となった. 各条件における AWWL の重み係数の平均値は, 視覚外条件において MD : 4.96 (SD = 1.17), PD : 2.46 (SD = 1.31), TD : 3.18 (SD = 1.44), OP : 2.89 (SD = 1.47), EF : 3.57 (SD = 1.35), FR : 1.89 (SD = 1.17) であり, 視覚内条件において MD : 4.75 (SD = 1.32), PD : 1.89 (SD = 1.22), TD : 3.03 (SD = 1.60), OP : 3.00 (SD = 1.63), EF : 4.14 (SD = 1.48), FR : 2.10 (SD = 1.13) となった.

以上の NASA-TLX における各単項目・RTLX・AWWL の結果は, 参加者が視覚内条件と比べて視覚外条件のときの方が負荷を大きく感じたことを示している.

二重課題法の結果

図 8 (左) は, 参加者のサブタスクの Precision, Recall, F-値の平均値を表している. Precision は Single または Triple の振動パターン発生時に参加者が正しく端末をタップできた割合を示し, Recall は Double または Quadruple の振動パターン発生時に端末をタップしない選択が正しくできた割合を表す. また F 値は Precision と Recall の調和平均を示している. Precision は t 検定の結果において, 視覚内条件の方が視覚外条件と比べて有意に高い結果となった ($t(27) = 3.31, p < 0.01$). 一方で Recall は両条件で有意差が生じなかった ($t(27) = 0.38, p = 0.70$). F-値は視覚内条件の方が視覚外条件と比べ有意に高い結果となった

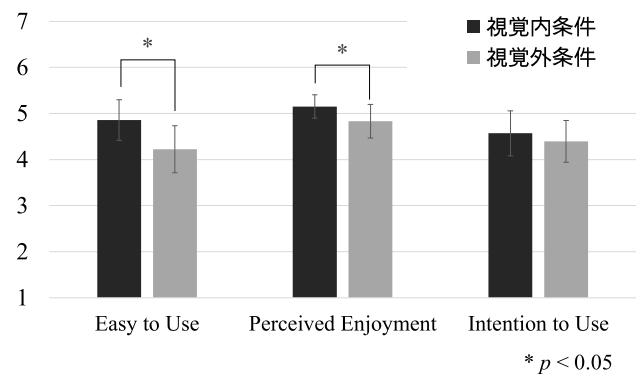


図 9 対話の印象評価のアンケートの平均値
 Fig. 9 Mean scores of questionnaire for evaluating conversational impression.

($t(27) = 2.88, p < 0.01$).

図 8 (右) は, サブタスクで参加者が正しいタップ操作をしたときの振動パターン発生後からタップまでの平均時間を表す. 本結果へ t 検定を実施したところ, 条件間の有意差は生じなかった ($t(27) = 0.20, p = 0.84$).

二重課題法の Precision と F-値の結果は, 参加者がサブタスク実施の精度において, 視覚内条件の方が視覚外条件よりも高いパフォーマンスを示したことを表している.

対話に対する印象評価

図 9 へ Easy to Use, Perceived Enjoyment, Intention to Use 項目の参加者回答の平均値を示す. t 検定の結果 Easy to Use ($t(27) = 2.37, p < 0.05$) と Perceived Enjoyment ($t(27) = 2.37, p < 0.05$) の値は, 視覚内条件が視覚外条件より有意に高い結果となった. 一方で Intention to Use の値では条件間で有意差は生じなかった ($t(27) = 0.96, p = 0.35$).

Easy to Use の結果は, 視覚内条件の方が視覚外条件と比べて参加者が簡単に 2 台ロボットと対話ができたと示している. また Perceived Enjoyment は, 視覚内条件の方が視覚外条件よりも, 参加者が対話の楽しさを強く感じたことを表している.

参加者の顔・肩方向の動作

対話中の参加者の顔および肩の水平方向の動きを下記に示す. 視覚内条件では 1 回の対話中に顔の動きが 15 deg/s の速度を超えたのは平均 1.42 回 (SD = 1.73) であり 30 deg/s を超える速度は一度も検出されなかった. また肩の速度は 15 deg/s の速度を超えた回数は平均 0.42 回 (SD = 0.63) であり 30 deg/s を超える速度は一度も検出されなかった.

一方, 視覚外条件では顔の動きが 15 deg/s の速度を超えたのは平均 35.83 回 (SD = 6.85) であり 30 deg/s を超える速度は平均 28.52 回 (SD = 5.69) であった. また肩の速度は 15 deg/s の速度を超えた回数は平均 17.83 回 (SD = 1.78) であり 30 deg/s を超える速度は平均 0.41 回 (SD = 0.79) であった.

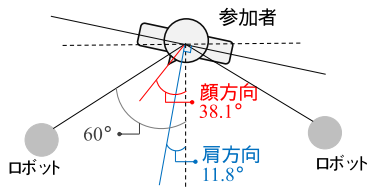


図 10 視覚外条件において参加者が各ロボットを見たときの顔および肩の角度の絶対値の最大値

Fig. 10 Mean absolute values of peak angles on participants face and shoulder when they looked at each robot in the outside visual field condition.

上記結果は、視界外条件では視界内条件と比べて顔および肩が水平方向により頻繁に素早く動いていたことを示す。

また視覚外条件において、参加者が視線を向けるロボットの対象を切り替えるごとに、参加者の顔および肩の正面方向からの最大角度を計測し、その最大角度の平均値を求めた。その結果、図 10 に示すように顔方向の正面からの平均角度は 38.1 度 (SD = 4.81) であり、肩方向の正面からの平均角度は 11.4 度 (SD = 3.75) であった、本結果より、参加者がロボットへ視線を向ける際に体よりも顔を主に向けており、顔方向の最大値はロボットの正面より 20 度ほど内側であることが分かる。

6. ケーススタディ

6.1 実験デザイン

ケーススタディでは比較実験の補足のため少数の高齢者を対象とした簡易なアンケートとインタビュー評価が中心に実施された。ケーススタディには 6 人の高齢者 (男性 3 人, 女性 3 人) が参加し、平均年齢は 75.5 歳 (SD = 5.68, 70~83 歳) であった。参加者のうち 3 名は要支援に認定されており、残りの 3 名は要介護に認定されていた。また全参加者は本実験がロボットとの初めての対話であった。

ケーススタディはトライアルと同じリハビリ施設にて実施され、利用された対話シナリオおよびロボットの動作・発話タイミングはすべて比較実験と同条件であった。ケーススタディでも参加者内実験が用いられ対話ロボットも同様の Sota が使用された。

対話後のアンケートにはトライアルと同じ内容の 4 つの単項目のアンケートが用いられた。ただし条件間でロボットの台数が変わるトライアルと異なり、ケーススタディでは両条件の区別が付きにくいいため、各質問にはロボットの位置に関する表現が加えられた (例: 2 台のロボットが前 (両脇) に居るときに、どのくらいロボット達と会話が弾んだか?)。

アンケートの回答には 7 段階のリッカート尺度 (1 = 低い, 7 = 高い) を用いた。またアンケート回答後は両条件のどちらをより好ましく感じたか、その理由についてのインタビューを行った。

6.2 結果

アンケートの平均値は、視覚内条件では Q1 (会話が弾んだか) : 5.16 (SD = 0.75), Q2 (会話の楽しさ) : 5.66 (SD = 1.51), Q3 (簡単に話せたか) : 5.66 (SD = 1.51), Q4 (また話したいと感じたか) : 6.66 (SD = 0.52) であった。また視覚外条件では Q1 : 4.83 (SD = 1.16), Q2 : 4.33 (SD = 1.21), Q3 : 4.83 (SD = 1.47), Q4 : 4.16 (SD = 1.60) であった。

インタビューでは、6 人中 5 人の高齢者が視覚外条件よりも視覚内条件の方が好ましいと回答し、残りの 1 名は両条件とも好ましいと回答した。またインタビューによる、視覚外条件が好ましくない理由についての回答は、大きく次の 2 つに分類することができた。

1 つ目は対話へ集中することの難しさである。この分類に属する回答は 6 人中 4 名の高齢者によってあげられた。本分類の代表的なコメントを次に紹介する。ID2: 「ロボットが両側にいるときは (視覚外条件)、体は疲れなかったんですけど、左右を見なきゃいけないから、その…気が散ってしまった感じがしましたね」

2 つ目は会話番を理解することの難しさである。本理由も 4 名の高齢者より回答された。代表的なコメントを 2 つ紹介する。ID1: 「はっきりとはいえないんですけど、両側にいるときは、前にいるときと比べて、話していいタイミングとか、いつ話してくるのかとか、分かりにくかったような気がしますね」、ID5: 「(視覚外条件では) ロボットが話しているのに気づかないことがあって、それで私が話そうとしたので、ロボット君が何ていっていたのか聞き取れないことがありました」。

7. 議論

比較実験において、客観的評価手法の NASA-TLX と、主観的評価手法の二重課題法の両方で、我々の仮説を支持する結果となった。二重課題法では、Precision と F 値において、視覚外条件よりも視覚内条件の方が、参加者のサブタスクへのパフォーマンスが高く、視覚内条件における対話時の負荷が視覚外条件におけるときより小さく、多くの認知資源をサブタスクに割り当てることができたと考えられるためである。また Easy to Use の結果からも、参加者が簡単に話をすることができたのは、視覚内条件の方が対話時の負荷が小さかったためと考えられる。

一方、NASA-TLX の作業達成度 (OP) や二重課題法の平均応答時間・Recall の値においては有意差が生じなかった。OP で有意差が生じなかった理由は、位置関係により対話負荷が増えた場合でも、その負荷が対話自体を失敗させるほど大きな影響力を持たなかったためと考えられる。また二重課題法の平均応答時間については、実験前のサブタスクの教示において参加者が過度にサブタスクに集中することで会話に影響を与えることを防ぐため、振動

パターンに対して可能な限り早く応答するよう指示しなかったことが有意差を生じなかった理由と考えられる。精度評価項目である Recall については、実験結果において Recall が Precision と比べて 20% も高い値となっていることから、参加者が振動パターンを分類できなかったときに、参加者は端末をタップするのではなく何もしないことを選択した可能性が考えられる。そのため誤ったタップの確率を表す Recall の値は両条件で違いが現れにくく有意差が生じなかったと考えられる。

Perceived Enjoyment は視界内条件の方が視覚外条件よりも有意に高い結果となっており、視覚外条件の高い対話負荷が会話の楽しさを下げたのが原因と考えられる。*Perceived Enjoyment* を構成する単体の質問項目では、「私はロボット達が魅力的だと感じた」、「私は面白いロボット達だと感じた」については、条件間での有意差は生じなかった。一方で「私はロボット達との会話を楽しむことができた」という項目については、視覚内条件の方が視覚外条件よりも有意に高い結果となった。以上の結果は、位置関係による対話負荷は会話に対する楽しさを下げたことがロボット自体の評価は下げないことを示唆している。Intention to Use の質問項目は、ロボット自体の使いやすさについて質問の焦点をあてており、使いやすさを問う対象が会話ではないことが有意差を生じなかった理由ではないかと考えられる。

ケーススタディにおいても、視覚内条件より視覚外条件の方が良いと回答した高齢者は 1 人もいなかった。比較実験によって示唆された不適切な位置関係における高い対話負荷および会話の楽しみの減少は、ケーススタディで高齢者が視覚内条件を好む原因となっておりトライアルにおいて 2 台ロボット条件が 1 台ロボット条件よりも低い評価となった原因の 1 つであると考えられる。

参加者の顔・肩の水平方向の動きは、視覚内条件よりも視覚外条件の方が多いことが示された。視覚外条件では参加者は発話中のロボットを見るために顔を動かしており、30 deg/s を超えた顔の動きの平均回数 28.52 はロボットからロボットへ発話番が移動した回数（前半シナリオ：27 回、後半シナリオ：30 回）に近い値となった。上記値は参加者が 1 分あたり 5 回程度の頻度で顔方向を変えたことを示している。視覚外条件で顔・肩の動きが多かったことは、身体的要求 (PD) の得点において視覚外条件が視覚内条件よりも有意に高い理由であると考えられる。

しかしながら、AWWL の評価において PD の係数は、視覚外条件において 2 番目に低く、視覚内条件において 1 番低い値となっている。そのため PD は全体の負荷において、支配的な要因ではないと考えられる。ケーススタディのインタビューにおいても、高齢者は身体的な問題よりも集中や対話番理解の難しさといった心理的な問題を指摘している。また比較実験における参加者のインタビューでも、身

体的な負担についての言及は、ほとんど確認されなかった (28 名中 1 名)。したがって、視覚外条件における位置関係の対話負荷は、身体的な負荷よりも、より心理的な負荷が主要因を占めていると考えることができる。

ケーススタディで言及されている視覚外条件での対話番の理解の難しさは、同様に比較実験の参加者からも、視覚外条件の方が視覚内条件よりも対話番の理解が難しいことがインタビューで言及された (28 名中 15 名)。視覚外条件において対話番が分かり難い理由の 1 つに、ロボットの視線情報が減ることが考えられる。たとえば、参加者 ID11 は「ロボットが横にいるときは (視覚外条件)、どちらのロボットが次話すのか良く分からなくて、次話していいのか迷った。ロボットが前にいるときは (視覚内条件) 2 人とも見えるので、横よりも (話してよい) タイミングが分かりやすかった。」と言及している。視覚外条件では、参加者は多くの場合において発話中のロボットを見ているため、もう一方のロボットの視線方向については同時に見ることはできない。そのため、参加者は 1 台のロボットの視線から次の話者が誰になるかを予想する必要があり、結果として対話番の分かりにくさにつながったと考えられる。

対話番の分かりにくさは、精神的要求 (MD) の項目以外にも、努力 (EF) や不安 (FR) といった心理的な負荷へ影響を与えたと考えられる。また次の対話番の予測を誤った場合、参加者は「ロボットの話を聞くのか」「ロボットへ話すのか」、話を聞く場合にはどちらのロボットが話すのかといった、対話状況を理解し直して、改めて集中する対象を決定する必要がある。そのため、視覚外条件では時間的圧迫感 (TD) についても有意に高い結果となったのではないかと考えられる。

8. リミテーション

本研究では、視覚内条件と視覚外条件の 2 つの位置関係を用いて、両条件で差が生じるかに焦点を当てて、対話中の負荷評価を実施した。今後は、視覚内条件と視覚外条件の間の角度条件においても調査を行い、対話の負荷が角度とともにどのように変化していくか、より詳細な評価を行う必要がある。

また比較実験とケーススタディの結果は、トライアルにおいて 2 台ロボット条件が 1 台ロボット条件よりも評価が低かった原因について 1 つの要因を示しているが、2 台ロボットの評価が低かった理由に他の原因が存在する可能性は否定できない。今後は、視覚内条件と同じ位置関係の 2 台ロボットと 1 台ロボットを比較した実験を行うことで、複数台ロボットを用いたコミュニケーションのメリットがどのような条件下で生じ、また失われてしまうかについては、さらなる調査が行われる必要がある。

9. 結論

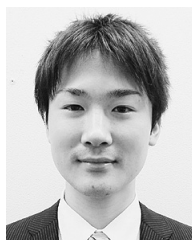
本論文では、人と複数ロボットの対話における、位置関係と対話負荷の関係について調査を行った。我々は、まず高齢者を対象としたトライアル調査の結果から位置関係と対話負荷についての仮説を立て、比較実験とケーススタディを通じて仮説検証を実施した。さらに本論文では、3つの調査で用いた、対話シナリオを元にジェスチャ・視線制御・対話タイミングを自動で生成する、複数ロボットの対話システムについて述べた。3つの調査を通じた、本研究の結論は以下の3つである。1: 対話中の位置関係において、対話の参加者が他参加者を同時に見ることができない場合、同時に見ることができる場合と比べ、参加者の受ける対話負荷は大きくなる。2: 対話中の位置関係によって生じる高い負荷は、人に対話自体をつまらなさと感じさせ、対話への印象を悪化させてしまう。3: 不適切な位置関係によって生じる負荷の要因は、顔や体の動きによる身体的な負荷だけではなく、むしろ対話に集中し辛い・対話番の認知の困難さによる精神的な要因が大きく寄与している。

本研究は、人と複数ロボットの対話において、位置関係と対話負荷に焦点を当てた最初の研究である。対話ロボットの開発者やアプリケーションデザイナー、特に高齢者など負荷に大きな影響を受ける相手を対象としている人達にとって、本研究が対話時の位置関係へより注意を払う動機付けとなれば幸いである。

参考文献

- [1] Iio, T., Yoshikawa, Y. and Ishiguro, H.: Pre-scheduled Turn-Taking between Robots to Make Conversation Coherent, *Proc. 4th International Conference on Human Agent Interaction, HAI*, pp.19–25 (2016).
- [2] Kanda, T., Ishiguro, H., Ono, T., Imai, M. and Mase, K.: Multi-robot Cooperation for Human-Robot Communication, *Proc. 11th IEEE International Workshop on Robot and Human Interactive Communication, RO-MAN*, pp.271–276 (2002).
- [3] Hall, E.T.: *The Hidden Dimension* (1966).
- [4] Kendon, A.: *Conducting interaction: Patterns of behavior in focused encounters*, Vol.7, Cambridge University Press (1990).
- [5] Kendon, A.: The Role of Visible Behavior in the Organization of Social Interaction, *Social Communication and Movement: Studies of Interaction and Expression in Man and Chimpanzee*, New York: Academic Press, pp.29–74 (1973).
- [6] Kuzuoka, H., Suzuki, Y. and Yamashita, J.: Reconfiguring Spatial Formation Arrangement by Robot Body Orientation, *Proc. 5th ACM/IEEE International Conference on Human-Robot Interaction, HRI*, pp.285–292 (2010).
- [7] Mumm, J. and Mutlu, B.: Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction, *Proc. 6th ACM/IEEE International Conference on Human-Robot Interaction, HRI*, pp.331–338 (2011).
- [8] Yamaoka, F., Kanda, T., Ishiguro, H. and Hagita, N.: How Close? A Model of Proximity Control for Information-presenting Robots, *Proc. 3rd ACM/IEEE International Conference on Human-Robot Interaction, HRI*, p.137 (2008).
- [9] Yamaoka, F., Kanda, T., Ishiguro, H. and Hagita, N.: Developing a Model of Robot Behavior to Identify and Appropriately Respond to Implicit Attention-Shifting, *Proc. 4th ACM/IEEE International Conference on Human-Robot Interaction, HRI*, pp.133–140 (2009).
- [10] Heerink, M., Kröse, B., Evers, V. and Wielinga, B.: The Influence of Social Presence on Acceptance of a Companion Robot by Older People, *Journal of Physical Agents*, Vol.2, No.2, pp.33–40 (2008).
- [11] Mitzner, T.L., Boron, J.B., Fausset, C.B., Adams, A.E., Charness, N., Czaja, S.J., Dijkstra, K., Fisk, A.D., Rogers, W.A. and Sharit, J.: Older Adults Talk Technology: Technology Usage and Attitudes, *Computers in Human Behavior*, Vol.26, No.6, pp.1710–1721 (2010).
- [12] Saaskilahti, K., Kangaskorte, R., Pieska, S., Jauhiainen, J. and Luimula, M.: Needs and User Acceptance of Older Adults for Mobile Service Robot, *Proc. 21st IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN*, pp.559–564 (2012).
- [13] Wolters, M., Georgila, K., Moore, J.D., Logie, R.H., MacPherson, S.E. and Watson, M.: Reducing working memory load in spoken dialogue systems, *Interacting with Computers*, Vol.21, No.4, pp.276–287 (2009).
- [14] Lohse, M., Rothuis, R., Gallego Pérez, J., Karreman, D.E. and Evers, V.: Robot Gestures Make Difficult Tasks Easier: The Impact of Gestures on Perceived Workload and Task Performance, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI*, pp.1459–1466 (2014).
- [15] Hayashi, K., Kanda, T., Miyashita, T., Ishiguro, H. and Hagita, N.: Robot Manzai – Robots’ conversation as a passive social medium, *Proc. 5th IEEE-RAS International Conference on Humanoid Robots, HUMANOIDS*, pp.456–462 (2005).
- [16] Vázquez, M., Steinfeld, A., Hudson, S.E. and Forlizzi, J.: Spatial and Other Social Engagement Cues in a Child-robot Interaction: Effects of a Sidekick, *Proc. 9th ACM/IEEE International Conference on Human-Robot Interaction, HRI*, pp.391–398 (2014).
- [17] Huettenrauch, H., Eklundh, K.S., Green, A. and Topp, E.A.: Investigating Spatial Relationships in Human-Robot Interaction, *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, pp.5052–5059 (2006).
- [18] Marshall, P., Rogers, Y. and Pantidi, N.: Using F-formations to Analyse Spatial Patterns of Interaction in Physical Environments, *Proc. ACM Conference on Computer Supported Cooperative Work, CSCW*, pp.3033–3042 (2011).
- [19] Setti, F., Russell, C., Bassetti, C. and Cristani, M.: F-Formation Detection: Individuating Free-Standing Conversational Groups in Images, *PLoS ONE*, pp.1–26 (2015).
- [20] Joosse, M.P., Poppe, R.W., Lohse, M. and Evers, V.: Cultural Differences in How an Engagement-seeking Robot Should Approach a Group of People, *Proc. 5th ACM International Conference on Collaboration Across Boundaries: Culture, Distance & Technology, CABS*, pp.121–130 (2014).

- [21] Patten, C.J.D., Kircher, A., Östlund, J. and Nilsson, L.: Using mobile telephones: Cognitive workload and attention resource allocation, *Accident Analysis and Prevention*, Vol.36, No.3, pp.341-350 (2004).
- [22] Kennington, C., Kousidis, S., Baumann, T., Buschmeier, H., Kopp, S. and Schlangen, D.: Better Driving and Recall When In-car Information Presentation Uses Situationally-Aware Incremental Speech Output Generation, *Proc. 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI*, pp.1-7 (2014).
- [23] Williams, K., Flores, J.A. and Peters, J.: Affective Robot Influence on Driver Adherence to Safety, Cognitive Load Reduction and Sociability, *Proc. 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI*, pp.1-8 (2014).
- [24] Yoshiike, Y., Silva, P.R.S.D. and Okada, M.: MAWARI: A Social Interface to Reduce the Workload of the Conversation, *Proc. 6th KES International Conference on Agents and Multi-Agent Systems, KES AMSTA*, pp.11-20 (2012).
- [25] Karatas, N., Yoshikawa, S., De Silva, P.R. and Okada, M.: NAMIDA: How to Reduce the Cognitive Workload of Driver, *Proc. 11th ACM/IEEE International Conference on Human-Robot Interaction, HRI*, pp.449-450 (2016).
- [26] Salthouse, T.A.: What and When of Cognitive Aging, *Current Directions in Psychological Science*, Vol.13, No.4, pp.140-144 (2004).
- [27] Häikiö, J., Wallin, A., Isomursu, M., Ailisto, H., Matinmikko, T. and Huomo, T.: Touch-Based User Interface for Elderly Users, *Proc. 9th International Conference on Human Computer Interaction with Mobile Devices and Services, MobileHCI*, pp.289-296 (2007).
- [28] Stephens, E.C., Carswell, C.M. and Mitzi, M. Schumacher: Evidence for an Elders' Advantage in the Naive Product Usability Judgments of Older and Younger Adults, *The Journal of the Human Factors and Ergonomics Society*, Vol.48, No.3, pp.422-433 (2006).
- [29] Mutlu, B., Forlizzi, J. and Hodgins, J.: A Storytelling Robot: Modeling and Evaluation of Human-Like Gaze Behavior, *Proc. 6th IEEE-RAS International Conference on Humanoid Robots, HUMANOIDS*, pp.518-523 (2006).
- [30] Hayashi, K., Sakamoto, D., Kanda, T., Shiomi, M., Koizumi, S., Ishiguro, H., Ogasawara, T. and Hagita, N.: Humanoid Robots as a Passive-Social Medium - A Field Experiment at a Train Station, *Proc. 2nd ACM/IEEE International Conference on Human-Robot Interaction, HRI*, pp.137-144 (2007).
- [31] Core, M.G. and Allen, J.F.: Coding Dialogs with the DAMSL Annotation Scheme, *Proc. AAAI Fall Symposium on Communicative Action in Humans and Machines*, pp.28-35 (1997).
- [32] Meguro, T., Higashinaka, R., Minami, Y. and Dohsaka, K.: Controlling Listening-oriented Dialogue using Partially Observable Markov Decision Processes, *Proc. 23rd International Conference on Computational Linguistics, COLING*, pp.761-769 (2010).
- [33] Kendon, A.: Some Functions of Gaze-direction in Social Interaction, *Acta Psychologica*, Vol.26, pp.22-63 (1967).
- [34] Imai, M., Kanda, T., Ono, T., Ishiguro, H. and Mase, K.: Robot Mediated Round Table: Analysis of the Effect of Robot's Gaze, pp.411-416 (2002).
- [35] Ishii, R., Otsuka, K., Kumano, S. and Yamato, J.: Prediction of Who Will Be the Next Speaker and When Using Gaze Behavior in Multiparty Meetings, *ACM Trans. Interactive Intelligent Systems*, Vol.6, No.1, pp.1-31 (2016).
- [36] Jovanovic, N., Akker, R.D. and Nijholt, A.: Addressee Identification in Face-to-Face Meetings, *Proc. 11th Conference of the European Chapter of the Association for Computational Linguistics, EACL*, pp.169-176 (2006).
- [37] Chen, L. and Harper, M.P.: Multimodal Floor Control Shift Detection, *Proc. 12th International Conference on Multimodal Interfaces, ICMI-MLMI*, pp.15-22 (2009).
- [38] Dielmann, A., Garau, G. and Bourlard, H.: Floor Holder Detection and End of Speaker Turn Prediction in Meetings, *Proc. 11th Annual Conference on the International Speech Communication Association, INTERSPEECH*, pp.2306-2309 (2010).
- [39] Jokinen, K., Furukawa, H., Nishida, M. and Yamamoto, S.: Gaze and Turn-Taking Behavior in Casual Conversational Interactions, *ACM Trans. Interactive Intelligent Systems*, Vol.3, No.2 (2013).
- [40] Seya, Y. and Watanabe, K.: Objective and Subjective Sizes of the Effective Visual Field during Game Playing Measured by the Gaze-contingent Window Method, *International Journal of Affective Engineering*, Vol.12, No.1, pp.11-19 (2013).
- [41] Hart, S.G. and Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, *Advances in Psychology*, Vol.52, pp.139-183 (1988).
- [42] De Waard, D.: *The Measurement of Drivers' Mental Workload*, Groningen University, Traffic Research Center Netherlands (1996).
- [43] Strayer, D.L. and Johnston, W.A.: Driven to Distraction: Dual-Task Studies of Simulated Driving and Conversing on a Cellular Telephone, *Psychological Science*, Vol.12, No.6, pp.462-466 (2001).
- [44] Kunar, M.A., Carter, R., Cohen, M. and Horowitz, T.S.: Telephone conversation impairs sustained visual attention via a central bottleneck, *Psychonomic Bulletin & Review*, Vol.15, No.6, pp.1135-1140 (2008).
- [45] Iwamura, Y., Shiomi, M., Kanda, T., Ishiguro, H. and Hagita, N.: Do Elderly People Prefer a Conversational Humanoid as a Shopping Assistant Partner in Supermarkets?, *Proc. 6th ACM/IEEE International Conference on Human-Robot Interaction, HRI*, pp.449-457 (2011).
- [46] Miyake, S. and Kumashiro, M.: Subjective mental workload assessment technique - An introduction to NASA-TLX and SWAT and a proposal of simple scoring methods, *The Japanese Journal of Ergonomics*, Vol.29, No.6, pp.399-408 (1993).



松元 崇裕

2012年慶應義塾大学情報工学科修士課程修了。同年日本電信電話株式会社入社，2016年慶應義塾大学後期博士課程入学，ヒューマンロボットインタラクションの研究に従事。現在，同大学後期博士課程学生。



後藤 充裕

2008年電気通信大学電気通信学部情報通信工学科卒業。2010年電気通信大学電気通信学専攻情報通信工学専攻修了。同年日本電信電話株式会社入社、ヒューマンコンピュータインタラクションの研究に従事。現在、NTT

サービスエボリューション研究所勤務。



石井 亮

2008年東京農工大学大学院工学府情報工学専攻修士課程修了。同年日本電信電話株式会社入社。現在、NTTメディアインテリジェンス研究所研究主任。2013年京都大学大学院情報学研究科博士後期課程修了。2011-2013年

成蹊大学客員研究員。ACM 15th Int. Conf. Multimodal Interaction (ICMI2014) Outstanding Paper Award 他受賞。電子情報通信学会，人工知能学会各会員。博士（情報学）。



渡部 智樹 （正会員）

1992年横浜国立大学工学部卒業。同年日本電信電話株式会社入社。主に、家電制御技術，生体状態計測・推定技術の研究開発に従事。現在、NTTサービスエボリューション研究所主任研究員。博士（工学）。電子情報通信学会

会員。



山田 智広 （正会員）

1992年新潟大学大学院修士課程修了。同年日本電信電話株式会社に入社。以来、デジタルコンテンツ配信，アプライアンス，ロボティクス技術に関する研究開発に従事。2007年MBA取得。2015年ネットワークドロボット

&ガジェットプロジェクトのプロジェクトマネージャー。2018年よりNTTサービスエボリューション研究所プロアクティブナビゲーションプロジェクトプロジェクトマネージャー。iRooBOネットワークフォーラム理事。電子情報通信学会 LOIS 研究会委員長，IEEE，ACM 各会員。



今井 倫太 （正会員）

1992年慶應義塾大学理工学部電気工学科卒業。1994年同大学大学院計算機科学専攻修士課程修了。同年NTTヒューマンインタフェース研究所入社。1997年ATR 知能映像通信研究所へ出向。2002年慶應大学大学院理工

学研究科後期博士課程修了。博士（工学）。2009～2010年シカゴ大学客員研究員。現在，慶應大学理工学部情報工学科教授およびATR 知能ロボティクス研究所研客員究員。2017年ドコモモバイルサイエンス賞社会科学部門優秀賞受賞。人型ロボットや自律エージェントと人とのインタラクションの研究に従事。電子情報通信学会，人工知能学会，日本認知科学会，日本ロボット学会，ヒューマンインタフェース学会，ACM，IEEE 各会員。