

複数ロボットの発話の重なりによって創発する空間の知覚

水丸 和樹^{1,a)} 坂本 大介¹ 小野 哲雄¹

受付日 2018年3月5日, 採録日 2018年9月7日

概要: 人間の集団には, 社会空間と呼ばれる特有の空間が形成されている. この空間は集団に属している人々が活発にコミュニケーションをとっているときに強く形成され, 集団に属していない第三者の行動はこの空間に影響される. また, 日常生活における人間同士の会話には無意識に発話の重なりが生じ, 対話相手への同調・関心・理解などを表現するとともに, 活発な会話を生み出す重要な要因となっている. 一方で, 近年, ヒューマノイドロボットが実用化されてきており, 実証実験がさかんに行われている. 将来, 人間が複数のロボットと共生する環境を想定するとき, ロボットの集団に形成される空間を考慮する必要があるにもかかわらずそのような研究は十分に行われていない. 本研究では, ロボットどうしの活発なコミュニケーションを発話の重なりによって実現し, それによって創発される空間を人間がどのように知覚するかについて調査した. 実験結果より, 発話の重なりは会話の活発さの印象を向上させ, ロボットの集団に創発される空間はその会話を観察した人間の行動に影響を与えるということが示された.

キーワード: ロボット, インタラクションデザイン, 発話の重なり, 社会空間, 行動分析

Perception of Space Emerged from Overlapping Speech by Multiple Robots

KAZUKI MIZUMARU^{1,a)} DAISUKE SAKAMOTO¹ TETSUO ONO¹

Received: March 5, 2018, Accepted: September 7, 2018

Abstract: A unique space called social space is formed in human groups. This space is strongly formed when people belonging to the group are actively communicating and will influence the behavior of a third party not belonging to the group. Moreover, overlapping speech occurs unconsciously in human everyday conversations, expressing entrainment, interest, understanding, and so on to the other party as well as becoming an important factor for producing active conversation. On the other hand, in recent years, humanoid robots have been put into practical use and demonstration experiments are actively being carried out. When assuming a future environment in which humans coexist with multiple robots, it is necessary to consider the space formed in a group of robots, but such research has not been conducted sufficiently. In this research, we implemented active communication between robots by overlapping their speech and investigated how humans perceived the space which emerged from it. As a result, it was indicated that overlapping speech improved the impression of conversation activity and that the space which emerged in the group of robots affected the behavior of the person observing the conversation.

Keywords: robot, interaction design, overlapping speech, social space, behavior analysis

1. はじめに

人々の集団には特有の心理的空間が形成されており, 社会空間と呼ばれている [1], [2]. 社会空間は集団に属してい

る人々が活発にコミュニケーションをとっており, 相互に作用しているときに強く形成される. 一方で, 集団に属していない第三者の行動はこの空間によって生じる境界線に影響される [3], [4], [5]. たとえば, 立ち話をしている集団を通り過ぎるとき, 集団の間を通らずに, 迂回するということがある. このように, 集団の影響によって行動を変化させる場面は, 日常生活の中で頻繁に発生する.

¹ 北海道大学
Hokkaido University, Sapporo, Hokkaido 080-8628, Japan
^{a)} mizumaru@complex.ist.hokudai.ac.jp

ある集団がどの程度活発にコミュニケーションをとっているか判断する要素として発話の重なりがある。日本語の会話分析において、会話の中で生じる発話の重なりは、対話相手への同調・関心・理解などを表しているということが主張されている [6], [7], [8], [9]。このような肯定的な考え方は日本語だけでなく英語の会話においても確認されている [10]。また、会話が活発になると発話の重なり頻度が多くなることも確認されており [11], [12], [13], 発話の重なりが活発なコミュニケーションを生み出す役割を果たしている。発話の重なりは、ロボットの対話設計では考慮されていなかったが、会話の活発さを高め、話者同士の連帯感を強める重要な要因となっている。

近年、ヒューマノイドロボットの実証実験がさかんに行われており、ロボットと共生する環境が構築されつつある。そのような環境では、複数体のロボットが存在し、ロボットからなる集団が構築される可能性がある。その場合、人の集団における社会空間と同様の影響がロボットの集団においても生じる可能性がある。本研究では、人が複数体のロボットの活発な会話を観察し、ロボットの集団によって構築される空間をどのように知覚するか明らかにすることを目的とする。これを明らかにすることで、複数体のロボットを用いて人の動線を制御するなどといった応用が考えられる。そのために、ロボットどうしの会話の活発さの印象変化と、それによる人の行動変容に注目した実験を行った。まず最初の実験では、発話の重なりと会話の活発さの印象との関係性を明らかにすることを目的として、クラウドソーシングを用いて、ロボット2体の発話の重なりによる印象の変化について調査した。その結果をふまえ、発話の重なりによってロボット間に創発される空間を人間がどのように知覚するか明らかにすることを目的として、ロボット2体の会話を1名の参加者が観察する実験を行った。本研究において、ロボットどうしの対話における発話の重なりとは、ある発話が終了する前にもう一方が発話を開始する状況で生じる重なりを指す。

2. 関連研究

2.1 社会空間

社会空間とは、いくつかの個人空間が融合して新たにできる心理的空間であり、人は他者が作る社会空間に影響を受けて行動する [1], [2]。この社会空間の影響を実験により明らかにしている研究がいくつかある。

Cheyne ら [3] は、2人からなる集団が廊下で立ち話をしている場合と話をせず階下を見下ろしているだけの場合において、2人の間を避けて通り抜ける歩行者の割合を調査した。その結果、立ち話をしている場合の方が2人の間を避けて通り抜ける割合が多かった。Lindskold ら [4] は、4人(男女2人ずつ)からなる集団が歩道に立っており、(1) Control: 4人が別々のことをしている、(2) Audience: 4

人がショーウィンドウを覗き込んでいる、(3) Calm Interaction: 静かに立ち話をしている、(4) Angry Interaction: 互いに怒っている、(5) Laughing Interaction: 互いに笑いあっている、という5つの条件で集団の間を通り過ぎる歩行者の割合を調査した。その結果、複数人がコミュニケーションをとっていると歩行者が判断した場合には、とっていないと判断した場合よりも集団の間を避けて通り抜ける割合が多かった。Knowles [5] は、大学の職員と学生が2人もしくは4人からなる集団を作って立ち話をしているときにその間を避けて通り抜ける歩行者の割合を調べた。その結果、地位が高いと判断された大学の職員の間を避けて通り抜ける歩行者の割合は多く、集団の人数が2人よりも4人の方が間を避けて通り抜ける割合が多いということも報告されている。これらの研究は集団に属する人々が相互に作用しており、活発にコミュニケーションをとっているときに社会空間が強く形成されるということを示している。

Imayoshi ら [14] は、人間の集団によって形成される社会空間を考慮したシステムとして人間の発話量を用いて会話の活性度を評価し、ユーザ間のスマートフォンの距離情報を用いて社会空間を認識するロボットシステムを提案した。実験では、2人の人間が会話しているときにロボットがその社会空間に介入し、ロボットに対する印象と影響について評価した。その結果、会話の活性度が高い場合よりも、低い場合の方が人間はロボットをより好意的に受け入れた。

以上のように、人間の集団による社会空間については研究されているが、ロボットの集団による空間については十分に研究が行われておらず、それによる影響も十分に考慮されていない。本研究では、ロボット2体が活発にコミュニケーションを行う際に、それを観察する第三者の人間がロボットの集団に創発される空間をどのように知覚するか注目した実験を行う。

2.2 発話の重なり

会話分析において、会話の中で生じる発話の重なりは、話者交代における誤りであったり、対話相手の発話権の侵害であるなど否定的にとらえられていた [15], [16]。しかし、次に示すいくつかの研究は日常会話には至るところで発話の重なりが生じており、肯定的にとらえられると指摘している。

藤井ら [6] は、日本人女性の友人同士の会話データを用いて、重なり頻度や機能などについて調査した。その結果、平均3.7発話に1度重なりが生じ、重なり全体の少なくとも40%以上は、先行発話への同意・共感・関心などを積極的に示すことで会話を促進させ、話者同士の連帯感を強める協力的な機能を持つことを示している。また、日本人の友人同士の雑談を分析した結果、対話相手の発話権を侵害するような発話の重なりは、わずか18.7%しか生じて

いなかったという報告もある [7]. 町田 [8] は, 初対面同士の会話データを分析し, 発話の重なりは話題となる内容が自分にとって身近であるということ素早く相手に伝え, 共感を深めるという機能を持つと主張している. この発話の重なりによる共感的な機能は日本語だけでなく英語の会話においても確認されている [10].

守屋ら [11] は, ビデオチャットシステムを用いて 3 人グループにおける会話音声の取得と会話活性度の主観的印象の評価を行い, それらの関係性を調査した. その結果, 会話活性度とオーバーラップ率 (発話の重なる頻度) には個人では中程度, グループ全体では強い相関があるということを示している.

ロボットどうしの対話設計に発話の重なりを取り入れる際, 複数ロボットによる効果についても考慮する必要がある. 神田ら [17] は, 2 体のロボットどうしが指差し対話を行う場合には指差しを行わずに対話する場合およびまったく対話を行わない場合と比べて, その様子を観察した人間は有意にロボットの発話内容を理解することを示した. さらに, ロボットに対してうなづきなどといった自然なコミュニケーションをとるということも示している. つまり, 複数ロボットによる発話と身体動作の相互作用は人とロボットが共生するうえで, 円滑なコミュニケーションを生み出すためには必要不可欠である.

以上のように, 人間の会話における発話の重なりについては研究されているが, ロボットの対話設計においては考慮されていない. また, 複数ロボットの対話を観察することの効果については検討されている. 本研究では, 発話の重なりを会話の活発さを測る指標に用いることを目的とした調査実験を行い, その後実際のロボットの対話を観察する実験を通して, 会話の活発さとロボットの集団に創発される空間の関係を調査する.

3. ロボット

本研究では, ソフトバンクロボティクス株式会社の人型ロボット「Pepper」を 2 体使用した. Pepper の基本仕様*1は, サイズ (高さ × 幅 × 奥行 = 1,210 × 480 × 425 [mm]), 重量 (29 kg), 自由度 (20) である. また, ロボットの発話とモーション作成には, Pepper の開発ツールである Choregraphe を使用した. Choregraphe では, 入力したテキストを自動で Pepper の発話に変換し, その発話に合わせたモーションを生成することができるため, それを利用してロボットの対話システムを実装した. これにより, 発話の開始と終了に合わせて身振り手振りなどの身体動作を付与することができる. 本研究はソフトバンクロボティクスの Pepper を活用し, 著者らが独自に実施したものである.

*1 <https://www.softbank.jp/robot/consumer/products/spec/>

4. 実験 1

本研究では, まずロボット 2 体の発話の重なりが, それを観察する人間にどのような印象を与えるかについて調べるために, クラウドソーシングを利用した実験を行った.

4.1 概要

実験 1 では, クラウドソーシングサービスである CrowdFlower *2 を用いて, ロボット 2 体の会話ビデオに関するアンケート調査を行った. CrowdFlower で集められた参加者は, Web ページで実験の説明を読んだ後, Google フォームによって作成されたアンケートページ (図 1 を参照) に誘導された. アンケートページでは, ロボット 2 体の会話についてのビデオを見た後, 以下の 5 段階形式 (5 が最も良い評価) のアンケートに回答した.

(1) How natural was the conversation of the robots?

(ロボットの会話はどの程度自然だったか?)

(2) How lively was the conversation of the robots?

(ロボットの会話はどの程度活発だったか?)

(3) How familiar were the robots with each other?

(ロボット 2 体の関係はどの程度親密だったか?)

アンケートではロボットどうしの会話の活発さだけではなく, 会話が自然に感じるかどうか, ロボット 2 体が親密かどうかについての印象を調査した. 会話が自然に感じるかどうかについては, ロボットの発話間隔を制御するうえで, 発話の重なりによって不自然な印象を与えるか調べる必要がある. さらに, 人間同士の会話において, 発話の重なりと親密さには高い相関がある [12] ことからロボットどうしの会話における親密さについても調査した.

図 1 アンケートページ (本研究はソフトバンクロボティクスの Pepper を活用し, 著者らが独自に実施したものである)

Fig. 1 Questionnaire pages (This research utilized Pepper of SoftBank Robotics and carried out independently by the authors).

*2 <https://www.crowdfunder.com/>

4.2 実験参加者

CrowdFlower を用いて集められた参加者の数は 1,233 名であり、日本人は含まれていなかった。アンケートでは、ボット (bot) などのランダムな回答をする参加者を不正な回答者と判断するなど、不正検知の仕組みを取り入れた。Koyama ら [18] はクラウドソーシングを用いた調査において、参加者に同じタスクを 2 回実行させ、2 回目の評価スライダーの終了を 1 回目とは逆に設定することで簡単な不正検知を行った。それを参考に、本実験では、アンケート内に質問項目 (3) と同じ質問を設定し、1 を最も良い評価とすることで不正検知を行った。たとえば、一方の評価で 4 を選んだにもかかわらず、もう一方で 3 を選んだ場合には評価が一致していないため、その参加者のデータは分析には利用しなかった。この結果、有効な回答者数は 721 名であった。

4.3 会話シナリオ

アンケートに使用するビデオでは、ロボット 2 体の発話間隔を以下の 3 つの条件として定義した。その定義をもとにロボットどうしの対話を実装し、その会話ビデオを撮影した。参加者はいずれかのビデオが含まれたアンケートページへランダムに誘導された。

- Overlap 条件
この条件ではロボットの発話が徐々に重なる。最初の発話間隔は 0 秒であるが、会話が進むにつれて徐々に発話が重なり、最後には 1 秒重なる (図 2 を参照)。会話ビデオ全体の長さは約 60 秒であった。また、この条件に割り振られた参加者数は 227 名であった。
- Default 条件
この条件ではロボットの発話間隔が常に 0 秒である (図 2 を参照)。つまり、ロボットの発話が終了するとすぐにもう一方のロボットの発話が始まる。会話ビ

デオ全体の長さは約 67 秒であった。また、この条件に割り振られた参加者数は 246 名であった。

- Delay 条件
この条件ではロボットの発話間隔が常に 1 秒である (図 2 を参照)。会話ビデオ全体の長さは約 79 秒であった。また、この条件に割り振られた参加者数は 248 名であった。

すべての条件における会話シナリオは同じである (表 1 を参照)。会話シナリオには公立高校の英語入試問題 (2002 年度 青森県) を使用し、会話の始まりと終わりの部分に変更を加えた (下線部)。この会話シナリオでは、それぞれのロボットが交互に発話し、発話回数はそれぞれ 7 回ずつであった。実験 1 で使用する CrowdFlower では英語が標

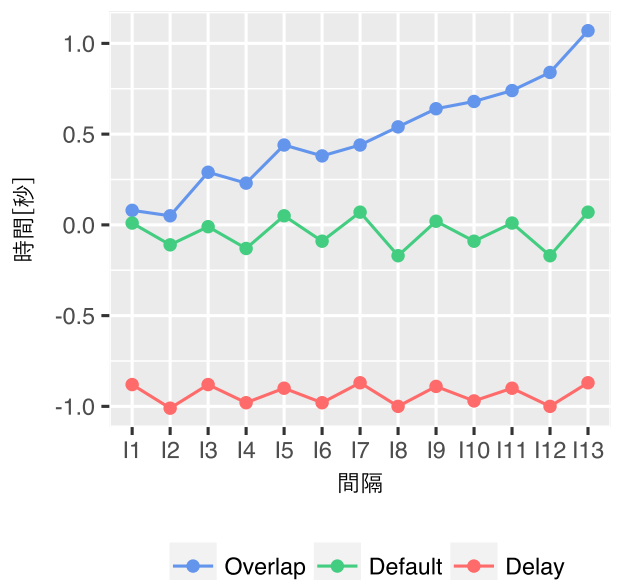


図 2 各条件のビデオにおける発話間隔 (実測値)
Fig. 2 Speech interval in the video of each condition (actually measured value).

表 1 会話シナリオ (R1 はロボット 1, R2 はロボット 2)

Table 1 Conversation scenario (R1 means Robot1, R2 means Robot2).

発話者	発話内容
R2	<u>Hi, Pepper.</u>
R1	<u>Hi.</u> Did you enjoy my English class yesterday?
R2	Yes, I did.
R1	I'm glad to hear that. Are you interested in studying English?
R2	Yes. I like to study English, but I have a problem.
R1	Oh, what is it?
R2	English is fun, but it is sometimes difficult for me. What is the best way to study English?
R1	If you use English every day, you will get better. Why don't you come to me and speak English?
R2	It's a great idea. I will try it.
R1	Good. And I would like to invite you to the English club. Many students come to the club and enjoy talking with me.
R2	Sounds great. I'd like to come. When should I visit the club?
R1	We have the club activities every Wednesday after school. Please come to the library.
R2	<u>Thank you very much.</u>
R1	<u>You're welcome.</u>

表 2 アンケート結果 (SD は標準偏差)

Table 2 Questionnaire results (SD means standard deviation).

	自然さ		活発さ		親密さ	
	平均	SD	平均	SD	平均	SD
Overlap	3.74	0.910	3.90	0.899	3.77	0.921
Default	3.77	0.973	3.79	0.897	3.65	0.990
Delay	3.73	0.933	3.69	0.954	3.77	0.957

表 3 多重比較の結果

Table 3 Results of multiple comparisons.

Overlap vs	自然さ	活発さ	親密さ
	<i>p</i> 値 (<i>r</i>)	<i>p</i> 値 (<i>r</i>)	<i>p</i> 値 (<i>r</i>)
Default	1.0 (0.02)	0.39 (0.06)	0.38 (0.06)
Delay	1.0 (0.01)	0.048 (0.1)	1.0 (0.003)

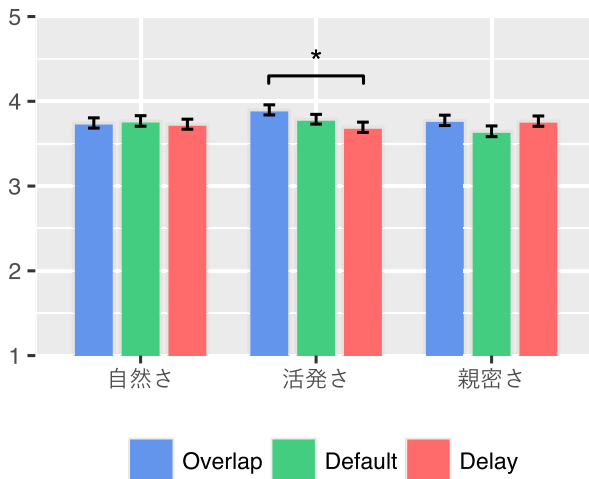


図 3 アンケート結果のグラフ (エラーバーは標準誤差)

Fig. 3 Graph of questionnaire results (error bar means standard error).

準言語であるが、参加者の中には、英語を母語としない参加者が含まれている可能性がある。実験 2 では、日本人参加者がロボットどうしの会話を観察する実験を行うが、実験 1 と条件を揃えるために、実験 1 と同じシナリオを使用する。また、実験 2 における参加者は大学生を想定している。これらより、実験 1 の日本人以外の参加者と実験 2 の日本人参加者の両方が理解しやすい英語の文章であると考えられ、交互に発話するというシナリオであることから、対話設計として発話の重なりを容易に実現できる公立高校の英語入試問題を利用した。

発話間隔 (図 2 を参照) は、「先発話者のロボットが発話の終了を検知した時刻」から「後発話者のロボットに発話を開始する命令を送信した時刻」を減算した。図 2 において、発話間隔が正の値のときは発話が重なることを表し、反対に負の値のときは発話が重ならず、小さくなるほど発話間隔が長くなることを表している。3つの条件 (Overlap, Default, Delay) の定義と図 2 の実測値は、システムの実行時間や通信時間などの影響から厳密には一致していない。

4.4 結果

実験 1 のアンケート結果とそのグラフを表 2, 図 3 にそれぞれ示す。アンケート結果について Mann-Whitney の U 検定によって Overlap 条件と Default 条件との比較、Overlap 条件と Delay 条件との比較を行った (Bonferroni

法による *p* 値調整)。その結果、Overlap 条件と Delay 条件の間には、アンケート項目 2 の「ロボットの会話はどの程度活発だったか」に有意な差が確認された ($p < .05$, $r = 0.1$) (表 3, 図 3 を参照)。一方で、Overlap 条件と Default 条件の間には、どのアンケート項目においても有意な差は確認されなかった (表 3, 図 3 を参照)。これらの結果は、発話間隔がつねに 1 秒である Delay 条件よりも、発話の重なりが生じる Overlap 条件の会話の方が活発に感じるということを示している。しかし、効果量 *r* が小さいことから、発話の重なりによる会話の活発さの印象を高める効果は小さいということが確認された。

5. 実験 2

実験 1 のアンケート結果から確認された発話の重なる効果をもとに、実験 2 では実際にロボット 2 体を用いて、ロボットどうしの会話を参加者が直接観察した。この実験を通して、参加者がロボットの発話の重なりによってロボットの集団に創発される空間をどのように知覚するか調査した。

5.1 概要

- 実験 2 では、ロボット 2 体と参加者 1 名による実験を行った (図 4 を参照)。実験の手順は以下のとおりである。
- (1) 参加者は実験の説明を受けた後、2体のロボットの前に立つように指示を受ける。
 - (2) 実験者の合図で参加者はロボットに挨拶をする。
 - (3) ロボットの会話が始まる。
 - (4) 実験者は別室で様子をモニタリングする。
 - (5) 会話の終わりに片方のロボットが参加者に机の上に置かれた教科書を見るようにすすめる。
 - (6) 実験者は会話終了後、およそ 30 秒後に実験室に入り、参加者に実験終了の旨を伝える。
 - (7) 参加者はアンケートに答える。

手順 (1) において、ロボットの会話が終了した後はその場に留まる必要はなく、自由に行動してもよいと教示した。一方で、会話が終了するまでは 2体のロボットの前にいるように教示した。さらに、手順 (6) において、誘導された場合はその時点で実験終了の旨を伝えた。実験終了時まで机の上に置かれた教科書に誘導されなかった場合は、実験者が参加者に対して教科書を見に行くように促して、どのような行動をとるか観察した。また、手順 (7) におい

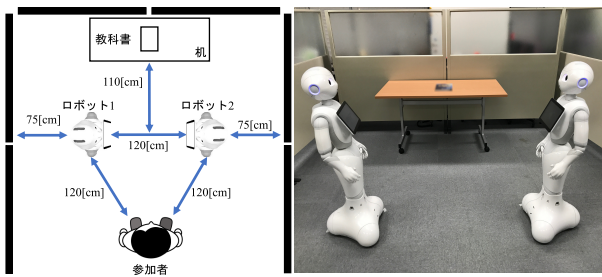


図 4 実験環境 (本研究はソフトバンクロボティクスの Pepper を活用し、著者らが独自に実施したものである)

Fig. 4 Experiment environment (This research utilized Pepper of SoftBank Robotics and carried out independently by the authors).

て、参加者は以下の 5 段階形式 (5 が最も良い評価) のアンケートに答えた。

- (1) ロボットの会話はどの程度自然だったか?
- (2) ロボットの会話はどの程度活発だったか?
- (3) ロボット 2 体の関係はどの程度親密だったか?

これらのアンケート内容は実験 1 のアンケートと同じ質問であり、実験に関する感想として自由記述欄を設けた。実験 2 において、ロボットの発話間隔の条件は実験 1 と同じく、Overlap 条件、Default 条件、Delay 条件の 3 条件とした。なお、それぞれの条件における定義は実験 1 と同じである。

5.2 実験参加者

参加者は日本人 39 名 (男性: 24 名, 女性: 15 名), 主に情報工学や薬学, 農学などを専攻する大学生・大学院生または高専・大学卒業の社会人であり, 年齢は 20~26 歳 (平均: 23.0, 標準偏差: 7.92) であった。また, 参加者の英語学習歴は最短でも 6 年以上であり, 「ロボット 2 体の英語の会話に参加する」という名目で実験に参加した。参加者はランダムに各条件 (Overlap, Default, Delay) に割り振られ, それぞれの条件で 13 名ずつ実験を行った (参加者間比較計画)。

5.3 会話シナリオ

実験 2 の会話シナリオでは, 参加者が会話に参加するために, 実験 1 で用いたシナリオの始めと終わりに以下の 2 つの変更を加えた。それぞれの変更部分を表 4 に示す。

- (1) 始めに参加者がロボットに挨拶をし, それぞれのロボットが順番に参加者に対して挨拶を返す (表 4 を参照)。
- (2) 終わりにロボット (ロボット 1) が机の上に置かれた教科書を参加者に対してすすめる (表 4, 図 4 を参照)。このとき, 参加者に机の上の教科書を意識させるために, 2 体のロボットが 1 度机の上の教科書の方を向いた後, 左側のロボットが参加者の方に頭を向けて, 発

表 4 会話シナリオの追加部分 (下線部)

Table 4 Additional parts of conversation scenario (underlined parts).

発話者	発話内容
参加者	(ロボットに挨拶をする, 例: Hi.)
R2	Hi. (参加者に挨拶する)
R1	Hi. (参加者に挨拶する)
R2	Hi, Pepper.
	...
R1	We have the club activities every Wednesday after school. Please come to the library.
R2	<u>OK. By the way, are there any textbooks you would recommended?</u>
R1	I recommend you the textbook on the table. (R1 が机の上に置かれた教科書を指し示す)
R1	<u>You should also look at it.</u> (R1 が参加者に向かっていう)
R2	Thank you very much.
R1	You're welcome.

表 5 アンケート結果 (SD は標準偏差)

Table 5 Questionnaire results (SD means standard deviation).

	自然さ		活発さ		親密さ	
	平均	SD	平均	SD	平均	SD
Overlap	4.31	0.855	4.54	0.519	3.77	1.01
Default	4.15	0.555	4.08	0.862	3.23	0.927
Delay	4.31	0.751	4.46	0.776	3.69	0.751

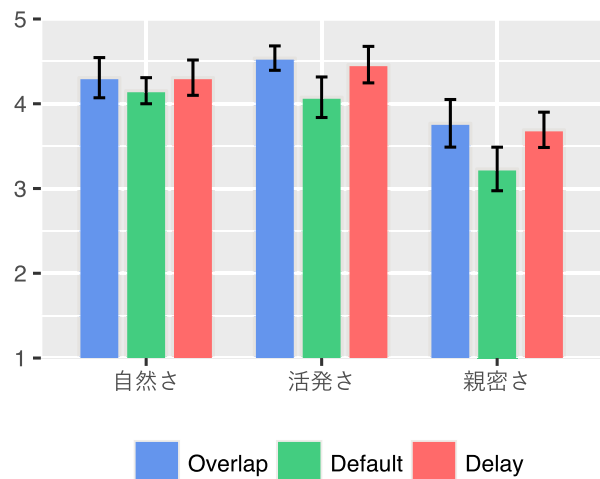


図 5 アンケート結果のグラフ (エラーバーは標準誤差)

Fig. 5 Graph of questionnaire results (error bar means standard error).

話した。

5.4 結果

実験 2 のアンケート結果とそのグラフを表 5, 図 5 にそれぞれ示す。アンケート結果について Mann-Whitney の U 検定によって Overlap 条件と Default 条件との比較, Overlap 条件と Delay 条件との比較を行った (Bonferroni

表 6 多重比較の結果

Table 6 Results of multiple comparisons.

Overlap vs	自然さ	活発さ	親密さ
	<i>p</i> 値 (<i>r</i>)	<i>p</i> 値 (<i>r</i>)	<i>p</i> 値 (<i>r</i>)
Default	0.73 (0.2)	0.42 (0.29)	0.52 (0.23)
Delay	1.0 (0.03)	1.0 (0.006)	1.0 (0.03)

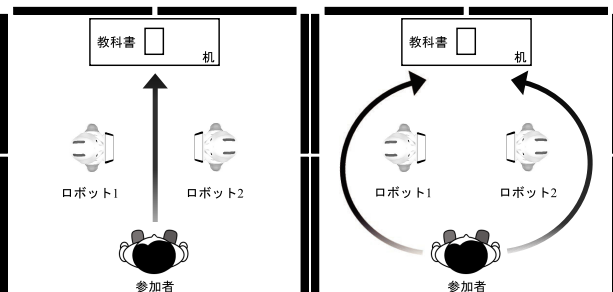


図 6 参加者の行動パターン

Fig. 6 Behavior patterns of participants.

表 7 各条件における参加者の行動パターンの集計結果

Table 7 Aggregate results of participants' behavior patterns in each condition.

	ロボットの間を通る	ロボットの間を避ける
Overlap	4	9
Default	7	5
Delay	9	2

法による *p* 値調整). その結果, Overlap 条件と Default 条件との間には, どのアンケート項目においても有意な差は確認されず, Overlap 条件と Delay 条件との間についても有意な差は確認されなかった (表 6, 図 5 を参照). これらより, 実験 2 において, ロボットの発話の重なりが参加者に与える印象 (自然さ・活発さ・親密さ) の効果は確認されなかった.

次に, 机の上に置かれた教科書を見に行く際に参加者がとった行動について分析する. 参加者の行動は以下の 2 つのパターンに分けられた (図 6 を参照).

- ロボット 2 体の間を通る.
- ロボット 2 体の間を避け, ロボットの後ろを通る.

また, それぞれのパターンの行動をとった参加者の人数を条件ごとに集計した (表 7 を参照). なお, 実験の説明では, 会話が終了した後に自由に行動してもよいと参加者に教示した. しかし, ロボットの会話シナリオが終了する前に教科書を見に行った参加者とシナリオが終了する前から見に行くような行動 (前に 1 歩踏み出すなど) をとり, 終了した直後に見に行った参加者がいた. それらの参加者は, ロボットの会話中であるにもかかわらず, ロボットの会話が終了したと判断していると考えられるため, 会話終了後の行動に注目する本研究では行動分析の対象から外した. 具体的には, Default 条件で 1 名, Delay 条件で 2 名除外した.

各条件における参加者の行動人数の結果についてフィッ

表 8 多重比較の結果

Table 8 Results of multiple comparisons.

Overlap vs	<i>p</i> 値 (ϕ)
Default	0.48 (0.277)
Delay	0.038 (0.51)

表 9 各条件における自由記述内容

Table 9 Contents of free description in each condition.

	アンケート内容 (同様の内容を記述した人数)
Overlap	動きがなめらかだった (1) 身振り手振りを使っていて驚いた (1) 会話のテンポが早かった (1) 最後の会話が早くなり聞きとりやすかった (1)
Default	自然に話していた (2) 会話がスムーズで動きも自然だった (2) 内容に沿った身振り手振りで, 違和感がなかった (1) 会話に不自然な間がなかった (1) 会話に混ざるタイミングをなかなか掴めなかった (1)
Delay	会話が自然だった (2) 会話がスムーズだった (1) 身振り手振りも人間のようになめらかだった (1) 会話間隔が少し長いように感じた (1) 思ったより会話に参加することはなかった (1)

シャーの正確確率検定 (Fisher's exact test) によって, Overlap 条件と Default 条件との比較, Overlap 条件と Delay 条件との比較を行った (Bonferroni 法による *p* 値調整). その結果, Overlap 条件と Delay 条件との間に有意な差が確認された ($p < .05, \phi = 0.51$) (表 8 を参照). 一方で, Overlap 条件と Default 条件との間には有意な差は確認されなかった. これらより, 本実験において発話間隔がつねに 1 秒である Delay 条件よりも, 発話の重なりが生じる Overlap 条件の方が, ロボットの間を避ける人が多く, 反対に Overlap 条件よりも Delay 条件の方が, ロボットの間を通る人が多いということが確認された.

5.5 アンケートの自由記述

アンケートの自由記述において, 参加者が記述した内容を抜粋し, 要約したものを表 9 に示す.

それぞれの条件において, ロボットの身振り手振りがあることによって, より人間に近い自然な会話を実現することができる可能性が示唆された. さらに, Overlap 条件では「会話のテンポが早かった」, Default 条件では「会話に不自然な間がなかった」, Delay 条件では「会話間隔が少し長いように感じた」とそれぞれ記述している参加者がいた. これより, それぞれの条件において参加者が主観的に感じる発話間隔は, 我々が定義した条件どおりであったことが示唆された.

6. 考察

6.1 発話の重なりの効果

実験1では、観察者から見て発話の重なりが会話の活発さの印象を向上させる効果があるということが確認された。統計分析に注目した際に、効果量が小さかった理由について考察する。守屋ら [11] は、参加者の興味などをもとに話題を設定することで会話活性度を制御していた。これに対して、実験1で使用したシナリオはすべての条件で統一されていた。さらに、クラウドソーシングで集められた参加者の中には、英語を母語としない参加者が含まれている可能性があることから、英語の理解度が低い参加者が混在しており、効果量が小さかった可能性がある。一方、実験2のアンケート結果では、会話の活発さの印象を向上させる効果があるということは確認されなかった。この結果について、実験参加者の属性の違いから考察する。実験1では、英語を標準言語とする CrowdFlower を利用して参加者が集められた。しかし、実験2では、日本人参加者のみで実験が行われた。それぞれの実験では同じ題材のシナリオを利用していたが、英語に対する理解度の差や、参加者の文化の違いが影響している可能性がある。また、実験2では実験1と比べて参加者の数が少なかったことから、発話の重なり効果が確認されなかった可能性がある。実験1と実験2において、観察者から見た「会話の自然さ」および「ロボットの親密さ」の印象の変化は確認されなかった。この理由についても、「会話の活発さ」と同様に、実験参加者の属性の違いなどが影響していると考えられる。

本研究では Overlap 条件において、ロボットの発話を徐々に重ねるように設計していた。これは、つねに発話を1秒重ねた場合にはロボットの会話のテンポが早く、参加者によるロボットの発話内容の聞き取りに影響があると考えたためである。実際、実験2のアンケートの自由記述(表9を参照)において「最後の会話が早くなり聞き取りにくかった」と記述している参加者がおり、徐々に発話を重ねる方がつねに重ねるよりも聞き取りやすく、つねに発話を1秒重ねると聞き取りが困難になると考えられる。

6.2 ロボットの集団に創発される空間

実験2では、それぞれの条件において、ロボットの集団に創発された空間に影響を受けて行動をした(ロボットの間を避けた)参加者がいた。一方で、発話の重なりによる会話の活発さの印象を向上させる効果が確認されなかったことから、本研究では会話の活発さとロボットの集団に創発される空間について関係性は明確にならなかった。これについては、実験2では参加者のタスクとして「教科書を見に行く」という行動をとることを意識させており、実際にはその途中の行動を観察していた。つまり、参加者にとっては無意識的にロボットの間を避けるという行動をとって

いた可能性があり、アンケートに答えるという意識的な行動とは異なると考えられる。また、ロボットの集団に創発された空間の影響は、先行研究 [3], [4], [5] で報告されている社会空間の影響と同じであることを示しており、ロボットの集団による影響が示唆された。その中でも、発話の重なりによって、より強い空間が形成されるということも示唆された。また、実験2では会話終了後の参加者の行動に注目していたが、先行研究 [3], [4], [5] では会話中の行動に注目している。したがって、ロボットどうしの会話中に人間がその集団に影響されるような行動をとるかどうかについても調査する必要がある。

6.3 複数ロボットの会話を観察する実験設定

実験2では、ロボットからの提案を会話の中で自然に行い、参加者の行動を促すことを意図していた。そのため、実験参加者を募集する段階、および実験における参加者への教示としては「ロボット2体の英語の会話に参加する」という名目であった。しかし、人がロボットに話しかけて応答するには設計されておらず、始めの挨拶だけという形になり、ロボットと人が会話する頻度は極端に少なくなっていた。実験2のアンケートの自由記述においても「会話に混ざるタイミングをなかなか掴めなかった」、「思ったより会話に参加することはなかった」と記述している参加者がおり、複数ロボットの会話を観察するという設定としては、人とロボットの関係性は多少曖昧であったと考えられる。

7. まとめ

本研究では、ロボットの対話設計における発話の重なり効果とロボットの集団に創発される空間の影響を調査した。実験1では、観察者から見て、発話の重なりは会話の活発さの印象を有意に向上させることが確認された。実験1の結果をふまえて実験2では、ロボットの集団に創発される空間において、人間の集団に特有の空間である社会空間と同様の影響があり、それを人間が知覚して行動が影響されることが確認された。これにより、ロボットどうしの会話において発話の重なりが生じている場合には、そうでない場合と比べて、観察者がロボットの集団に創発される空間を侵害しないような行動をとることが考えられる。つまり、複数ロボットの相互作用によって、人はそれらの集団を尊重するような行動をとり、本研究はその基礎的な知見を与えるものである。また、人とロボットの対話においても発話の重なりという対話設計を組み込むといった応用や、複数体のロボットによって人の動線を制御するような空間を形成することができると考える。本研究では、ロボットどうしの会話についてのアンケートは3項目(自然さ、活発さ、親密さ)のみであったが、さらに発話の重なりによる効果を調査していく必要がある。また、多く

の参加者に対して、ロボットの集団による影響を調査していくことも今後の展望としてあげられる。

謝辞 本研究は、科学研究費補助金・新学術領域研究「認知的インタラクシオンデザイン：意思疎通のモデル論的理解と人工物設計への応用」（領域：4601）からの支援を受けて実施された。

参考文献

- [1] 渋谷昌三：社会空間の基礎的研究，心理学研究，Vol.47, No.3, pp.119–128 (1976).
- [2] 渋谷昌三：人と人との快適距離—パーソナルスペースとは何か，日本放送出版協会 (1990).
- [3] Cheyne, J.A. and Efran, M.G.: The Effect of Spatial and Interpersonal Variables on the Invasion of Group Controlled Territories, *Sociometry*, Vol.35, No.3, pp.477–489 (1972).
- [4] Lindsfold, S., Albert, K.P., Baer, R. and Moore, W.C.: Territorial Boundaries of Interacting Groups and Passive Audiences, *Sociometry*, Vol.39, No.1, pp.71–76 (1976).
- [5] Knowles, E.S.: Boundaries around group interaction: The effect of group size and member status on boundary permeability, *Journal of Personality and Social Psychology*, Vol.26, pp.327–331 (1973).
- [6] 藤井桂子，大塚純子：会話における発話の重なりについて：協力的側面を中心に，言語文化と日本語教育，Vol.8, pp.1–13 (1994).
- [7] 藤井桂子：発話の重なりについて：分類の試み，言語文化と日本語教育，Vol.10, pp.13–32 (1995).
- [8] 町田佳世子：初対面の会話における発話の重なりの効果，北海道東海大学紀要．人文社会科学系，Vol.15, pp.189–210 (2002).
- [9] 生駒幸子：日常会話における発話の重なり機能，世界の日本語教育．日本語教育論集，Vol.6, pp.185–200 (1996).
- [10] Tannen, D.: *Conversational Style: Analyzing Talk Among Friends*, Ablex Publishing (1984).
- [11] 守屋悠里英，田中貴紘，宮島俊光，藤田欣也：ボイスチャット中の音声情報に基づく会話活性度推定方法の検討，ヒューマンインタフェース学会論文誌，Vol.14, No.3, pp.283–292 (2012).
- [12] 西村良太，北岡教英，中川聖一：音声対話における韻律変化をもたらす要因分析（〈特集〉リズムとタイミング），音声研究，Vol.13, No.3, pp.66–84 (2009).
- [13] 豊田 薫，宮越喜浩，山西良典，加藤昇平：発話状態時間長に着目した対話雰囲気推定，人工知能学会論文誌，Vol.27, No.2, pp.16–21 (2012).
- [14] Imayoshi, A., Munekata, N. and Ono, T.: Robots That Can Feel the Mood: Context-aware Behaviors in Accordance with the Activity of Communications, *Proc. 8th ACM/IEEE International Conference on Human-robot Interaction, HRI '13*, pp.143–144, IEEE Press (2013).
- [15] Sacks, H., Schegloff, E.A. and Jefferson, G.: A Simple Systematics for the Organization of Turn-Taking for Conversation, *Language*, Vol.50, No.4, pp.696–735 (1974).
- [16] Zimmerman, D.H. and West, C.: Sex roles, interruptions and silences in conversation, *Language and Sex: Difference and Dominance*, Thorned, B. and Henley, N. (Eds.), pp.105–129, Newbury House (1975).
- [17] 神田崇行，石黒 浩，小野哲雄，今井倫太，中津良平：人-ロボットの対話におけるロボット同士の対話観察の効果，電子情報通信学会論文誌 D-I, Vol.J85-D-I, No.4, pp.380–389 (2002).

- [18] Koyama, Y., Sato, I., Sakamoto, D. and Igarashi, T.: Sequential Line Search for Efficient Visual Design Optimization by Crowds, *ACM Trans. Graph.*, Vol.36, No.4, pp.48:1–48:11 (2017).



水丸 和樹

1995年生。2018年北海道大学工学部情報エレクトロニクス学科卒業。同年同大学大学院修士課程進学。人とロボットの関わりについて興味を持つ。



坂本 大介（正会員）

2008年公立ほこだて未来大学大学院システム情報科学研究科博士（後期）課程修了。博士（システム情報科学）。国際電気通信基礎技術研究所（ATR）でインターン，東京大学で日本学術振興会特別研究員 PD，JST ERATO 五十嵐デザインインタフェースプロジェクト研究員，東京大学大学院情報理工学系研究科コンピュータ科学専攻助教，特任講師を経て現在，北海道大学大学院情報科学研究科准教授。人とロボットを含む情報環境とのインタラクシオン設計に関する研究に従事。ACM 会員。



小野 哲雄（正会員）

1997年北陸先端科学技術大学院大学情報科学研究科博士後期課程修了。同年（株）ATR 知能映像通信研究所客員研究員。2001年公立ほこだて未来大学情報アーキテクチャ学科助教授，2005年同学科教授。2009年北海道大学大学院情報科学研究科教授，現在に至る。博士（情報科学）。ヒューマンエージェント/ロボットインタラクシオン（HAI/HRI），インタラクティブシステムに関する研究に従事。電子情報通信学会，ヒューマンインタフェース学会，認知科学会，ACM 各会員。