

対話意欲変動要因としての 対話ロボットの複数台化とジェスチャ表出の有効性評価

Effective evaluation of multiple robots and gestures
as dialogue motivation factor

丸山 洸太¹ 大和 淳司¹ 杉山 弘晃²

Kota MARUYAMA¹, Junji YAMATO^{1,2}, and Hiroaki SUGIYAMA²

¹ 工学院大学情報学部

¹ Faculty of Informatics, Kogakuin University

² NTT コミュニケーション科学基礎研究所

² NTT Communication Science Laboratories

Abstract: In this paper, we analyzed influence of multiple robots and gestures on dialogue motivation by making user interact with dialog robots. As a result, dialogue motivation increase for two robots rather than one robot, and rather dialog with gesture than dialog without gesture.

1. はじめに

近年, Apple の Siri や Softbank の Pepper など, 様々な対話システムが普及している. Siri などに見られる対話システムはアプリ内で完結しており, その手軽さから多くのユーザが利用している. しかし人同士の対話では, スキンシップやジェスチャなど身体を用いた非言語情報の共有が多く見られる. そのため, より自然な対話を実現するために, 身体性を持つ対話ロボットを用いた研究が行われている. 例えば, 高齢者の孤独死や認知症を防ぐために, 対話ロボットとのコミュニケーションを利用する研究が行われている[9]. このように, これからはユーザと雑談することを目的とした雑談対話システムの需要が増していくと予測される.

しかし現状の対話システムは, ユーザの話す話題に対応できなかつたり, 文脈保持がされず対話破綻を引き起こすなど, 課題も多い. このような課題はユーザの対話意欲を低下させるだけでなく, 「ユーザと対話する」という対話システム本来の目的も達成されなくなってしまう.

2. 関連研究

現在の対話システムは一問一答形式の対話であれば違和感を感じないような返答が可能になってきた. しかし対話を続けていくにつれて, 直前に話した文脈が保持されず, 対話破綻をきたすケースがある.

対話破綻はユーザの対話意欲を削ぐだけでなく, 対話が続けられなくなる場合が多いため, 対話破綻は解決すべき課題である. 杉山ら[3]は対話ロボットの台数を1台から2台にすることで, 対話破綻感が軽減することを示した. また, 北條ら[7]は対話システムとのインターフェースに着目し, テキスト, 音声, 実機対話の3つのメディアのうち, 実機対話のが最も対話破綻検出率が低かったと報告している.

対話ロボットと対話する上で, 「ロボットと意思疎通が取れている・対話ができている」といった対話感をユーザに与えることは重要である. 飯尾ら[2]は1台の対話ロボットと対話するよりも複数台の対話ロボットと対話する方が対話感を多く与えられたことを報告している.

これらの研究は, 対話ロボットの複数台化がユーザにとってポジティブな影響を与えていることを示唆している. しかし条件によってはロボットを複数台化するとネガティブな影響を与える可能性もある. 例えば, 非言語情報の共有がなければユーザが複数ロボットのうちどの個体が発話者か, 発話者の想定する次話者が誰かを都度確実に認識することは実は容易ではない. そのため, ロボット発話が終わったと思いうるユーザが発話しだしたら, 実はもう一台のロボットが発話するターンだった, という状況が発生し得る. だがこのような状況は, ジェスチャなどの非言語情報を適切に扱うことで回避できる可能性がある.

そこで本研究では, まず対話ロボットを1台から

2 台を増やすことで、ユーザの対話意欲が向上するか検証を行う。次に、ジェスチャ表出が対話意欲に影響するかについても分析を行う。

3. 対話実験

3.1. 実験目的

ユーザと対話ロボットを対話させたとき、ロボット1台との対話よりも2台との対話の方が対話意欲が高くなることを検証する。また、ジェスチャ表出を行わないロボットとの対話よりも、ジェスチャ表出を行うロボットとの対話の方が対話意欲が高くなることを検証する。

3.2. 評価指標

本実験で用いる評価指標を表1にまとめた。ここでは対話実験中にユーザの入力した「もっと話したい」と「もう話したくない」の評価のうち、「もっと話したい」の割合を対話意欲とする。

また、ストレス指標として有効性が示されている唾液アミラーゼ活性（以降 sAA：salivary Amylase Activity と呼ぶ）[6]を対話前後のストレス変化の指標として用いた。唾液アミラーゼ活性とは、唾液中に含まれるアミラーゼの触媒能力の尺度のことで、単位は主に気温や湿度を考慮した Unit が用いられる。文献[6]によれば、暗算後に計測した sAA と心拍・スキンコンダクタンス・緊張に対する主観評価（VAS：Visual Analogue Scale）のいずれの項目も有意に上昇し、sAA とそれぞれの評価項目間に有意に相関が見られたとの報告があり、ストレス増大が sAA の上昇につながる事がわかっている。ただし sAA の値そのものは個人差の影響が大きいため、対話前後の sAA 比の値を測ることとした。sAA 比から、各対話でどれだけ sAA が変化したかがわかる。

また、ユーザが発した1発話の平均文字数を発話長とした。発話長が長ければ、その対話において多く発話したものとする。

表 1. 実験で用いた評価指標

評価指標	評価指標の計測方法
対話意欲	「もっと話したい」「もう話したくない」のうち、「もっと話したい」の割合
sAA	対話前後で計測する唾液アミラーゼ活性の値
sAA 比	対話後 sAA/対話前の sAA
発話長	1 発話にユーザが発した平均発話文字数

3.3. 実験環境

本実験では以下の機器を用いて対話実験を行なう。

- 対話ロボット Sota
- PC (Windows10)
- モニタ
- 指向性マイク
- スピーカ
- USB オーディオ
- Raspberry Pi 3B
- ボタン
- レバー
- 唾液アミラーゼモニター
- 唾液採取チップ

対話実験を行なった実験室の概観図を図1に示す。残響の少ない防音室で対話実験を行なった。実験者用の机と実験協力者用の机を用意した。実験協力者用の机にコントローラ（Raspberry Pi に python で実装）を入れた箱を設置し、その上に音源分離機能を有する高指向性マイクをユーザに向けて設置した。箱には対話意欲を評価するボタンと、質問したロボットがどちらかを入力するレバーを取り付けた。実験協力者側から見て箱の奥側に対話ロボットを設置し、さらにその奥にスピーカを設置した。

実験者用の机には、対話システムを構成する PC を置いた。また、コントローラへ入力信号を送るためのキーボード、マウスと、コントローラからの出力を表示するためのモニタを設置した。

本実験は実際にユーザが自然に対話する状況を想定しているため、通常の対話シーンではユーザが見るはずのない配線や PC などを見せないよう、実験者用の机と実験協力者用の机の間にパーティションを設置した。

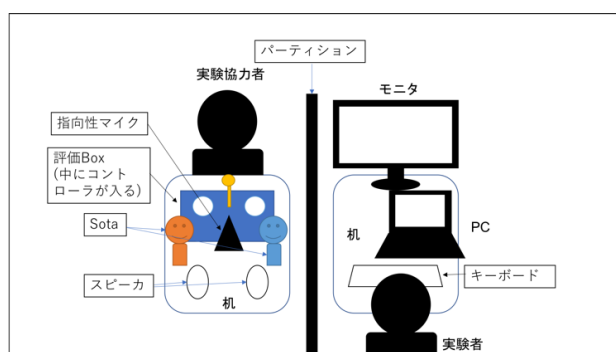


図 1. 実験室の概観図

3.4. 実験条件

本実験では対話ロボットの台数とジェスチャの有無を変化させて条件間の比較を行う。表 2 のように、実験協力者を 1 台のロボットと対話させる条件(以降条件 1 と呼ぶ)、2 台のロボットと対話させる条件(条件 2)、ジェスチャがない 2 台のロボットと対話する条件(条件 3)に分けて対話実験を行なった。

実験協力者は日本語話者 17 名(男 11 名, 女 6 名, 20-23 歳, 口腔内疾患が認められない人)で実験を行なう。各条件における実験協力者数は条件 1 が 4 人, 条件 2 が 9 人, 条件 3 が 4 人であった。本実験は被験者間実験のため、実験協力者は単一の条件でのみ対話実験を行う。

表 2. 実験条件

		
条件1	条件2	条件3
1台対話	2台対話	2台対話
ジェスチャ有	ジェスチャ有	ジェスチャ無

3.5. 仮説

本実験の仮説として、以下の仮説 1 から仮説 4 までの仮説を立てた。特に、仮説 3 と仮説 4 は以下の表 3 のような関係になると考えた。

- 仮説1. 条件 1 よりも条件 2 の方が対話意欲が高くなる
- 仮説2. 条件 3 よりも条件 2 の方が対話意欲が高くなる
- 仮説3. 対話意欲が高い状況では、発話長が短いグループ(以降グループ 2 と呼ぶ)よりも発話長が長いグループ(以降グループ 1 と呼ぶ)の方が sAA 比が小さい
- 仮説4. 対話意欲が低い状況では、発話長が長いグループ(以降グループ 3 と呼ぶ)よりも発話長が短いグループ(以降グループ 4 と呼ぶ)の方が sAA 比が小さい

表 3. 発話長と sAA 比の関係 (仮説)

	発話長が長い	発話長が短い
対話意欲が高い	グループ 1 sAA 比 小	グループ 2 sAA 比 大
対話意欲が低い	グループ 3 sAA 比 大	グループ 4 sAA 比 小

3.6. 実験手順

実験手順のフローチャートを図 2 に、実験中の様子を図 3 に示す。

実験協力者に対話ロボットと対話することを教示する。教示段階で口腔内に異常がないか、実験開始 1 時間前から食事を摂取していないか確認する。その後 sAA を計測し、計測が終わってから実験協力者を 5 分間安静にさせ、5 分経過した後に再び sAA を計測する。計測が終わったら、対話中の対話意欲の評価と、質問したロボット推定に使うレバー操作に十分慣れるまで訓練を行う。訓練が終了したら対話実験を開始する。

対話実験中はユーザ発話前に毎回、「もっと話したい」か「もう話したくない」かの評価を手元のボタンで入力する。その後、条件 2 と条件 3 の場合は、最後に質問したロボットがユーザから見て右側か左側かを、手元のレバーを傾けることで入力させる。条件 1 の場合はそのまま次の発話に移る。

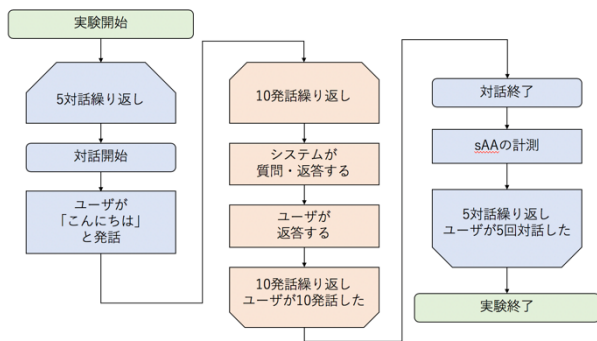
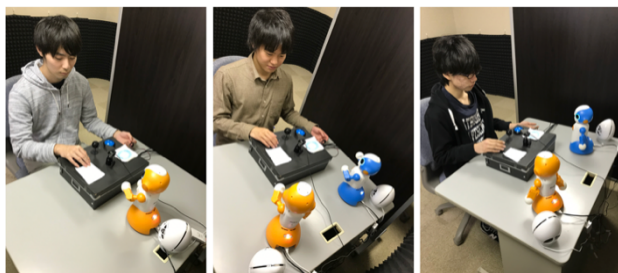


図 2. 実験手順フローチャート



(a). 条件 1 (b). 条件 2 (c). 条件 3
図 3. 実験時の様子

4. 実験結果

本実験で得られた対話意欲の結果を図4と図5に示す。

4.1. 条件1と条件2の比較結果

対話実験の結果、条件1よりも条件2の方が対話意欲が高くなる結果が得られた。条件1では79.9%の評価が「もっと話したい」だったのに対して、条件2は89.1%の評価が「もっと話したい」となった。この2条件間で χ^2 検定を行なった結果、 $p < 0.01$ となり有意差が認められた。

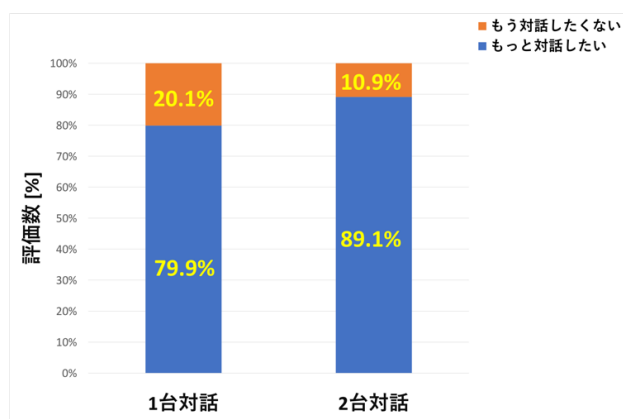


図 4. 条件1と条件2の対話意欲

4.2. 条件2と条件3の比較結果

対話実験の結果、条件3よりも条件2の方が対話意欲が高まる結果が得られた。条件3での評価のうち、68.3%が「もっと話したい」だった。この2条件間で χ^2 検定を行なった結果、 $p < 0.01$ となり有意差が認められた。

また、2台あるロボットのうち、質問を投げかけたロボットを推定した結果、条件2で派全ての発話に対してどちらのロボットが質問したか正しく認識できていたのに対して、条件3では全ての発話に対して正しく認識できていないことがわかった。

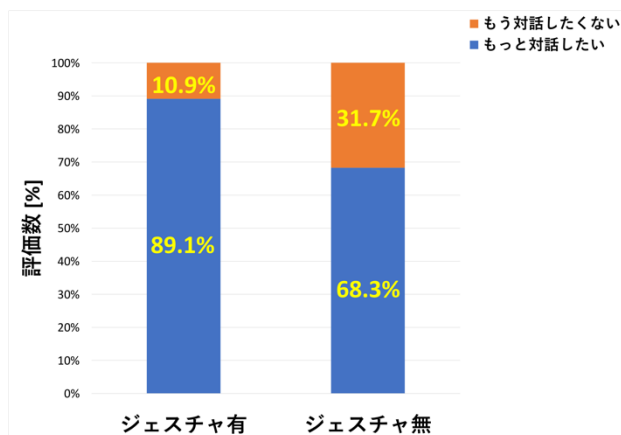


図 5. 条件2と条件3の対話意欲

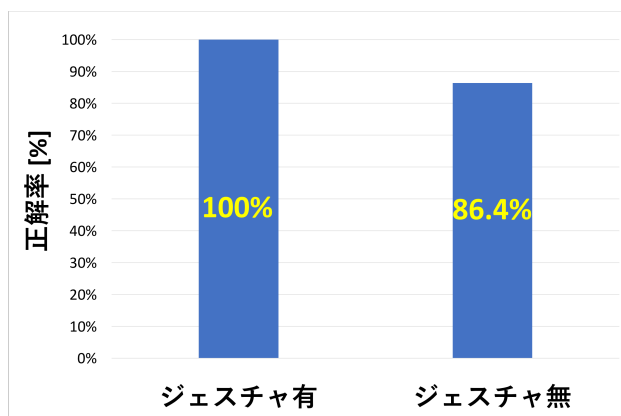


図 6. 発話ロボットの推定結果

4.3. 発話長と sAA 比の関係

発話長と sAA 比の関係を表4に示した。この結果から、対話意欲が高い場合は、発話長が短い対話よりも発話長が長い対話の方が sAA 比が小さくなることがわかった。また、対話意欲が低い場合は、発話長が長い対話よりも発話長が短い対話の方が sAA 比が小さくなることがわかった。

表 4. 発話長と sAA 比の関係 (結果)

	発話長が長い	発話長が短い
対話意欲が高い	1.29	1.32
対話意欲が低い	1.14	1.06

5. 考察

5.1. ロボット台数増加に伴う要因

条件1よりも条件2の方が対話意欲が高くなった。これは1台対話よりも2台対話の方がユーザへの負担が小さかった影響があると考えられる。条件1の場合は1対話に占めるユーザ発話の割合が、2台対話の時と比べて相対的に大きくなる。2台対話だとロボット同士の対話を挟むが、1台対話だとそのような間がなく、すぐにユーザへ質問を投げかける。そのためユーザは質問せめられているという感覚を抱きやすく、対話意欲の低下につながったと考えられる。また、実験協力者がロボットの質問に答える前に、ロボット同士の対話を聞く機会があったことも負担軽減要因としてあげられる。ロボット同士の対話は、その後ユーザへ尋ねる質問内容に関連することが多いため、ユーザは対話を聞くことで余裕を持って返答することができる。1台対話だと質問された直後にすぐ返答しないと沈黙になるため、心理的負荷が2台対話に比べて大きいものと考えられる。

5.2. ジェスチャ有無に伴う要因

条件2と条件3を比較すると、条件2の方が対話意欲が高い結果となった。これは2台対話の中でも、ターンテイキングが円滑に行われやすい条件と、そうでない条件だったためである。条件2の場合、ロボットの顔の向きから、もう一台のロボットに向けて話しているのか、ユーザに向けて話しているのが視覚的に判断できる。しかし条件3だとジェスチャがないため視覚的に判断できない。そのためロボットの質問が誰に向けられたものなのかが判断できず、ユーザ発話とシステム発話が重なるという状況が発生した。このような発話の重なりはユーザにネガティブな感情を起しやすく、対話意欲に差が出た要因であると考えられる。

本実験から、ロボットを複数台化することに伴う環境の変化は、ユーザに対してポジティブな影響だけでなくネガティブな影響も与え得ることが明らかとなった。条件3の対話意欲は条件2はもとより、条件1よりも低い結果であることがその一例と言える。この問題に対しては、発話時のジェスチャ表出を取り入れることで、発話ロボットが話しかける相手が誰かは明示することができる。しかし問題はターンテイキングにおける次話者の指定にどどまらない。図6から分かるように、ジェスチャなし条件(条件3)においては、発話しているロボットがどちらな

のかすらユーザは正しく判断できないケースが生じていた。

これらの分析を踏まえると、対話ロボットを複数台化することで発生する課題は、ユーザが発話者および期待される次発話者を正しく認識できない状況が発生するという点である。そしてそれはジェスチャなどの非言語情報を含めたロボットの発話行動を適切にデザインすることで解決する必要があると考えられる。

5.3. 発話長と sAA 比の関係

対話意欲が高いグループ1とグループ2の対話は、ユーザがたくさん話したいと感じていた対話だと言える。そのため、発話長が長いグループ1は欲求が満たされた状態であり、グループ2は欲求が満たされなかったグループだと言える。そのためグループ1よりもグループ2の方がsAA比が大きくなったと考えられる。

対話意欲が低いグループ3とグループ4の対話は、ユーザはあまり話したくない対話だったと言える。発話長が短いグループ4は話したくないから話さなかったグループであり、対話意欲が低いにもかかわらず発話長が長かったグループ3に比べてストレスが溜まりにくいグループだったと考えられる。これらの知見を積み上げ、対話意欲や対話の快適性の客観計測を目指していきたい。

6. まとめ

本研究では、対話ロボットの複数台化とジェスチャ表出の有無を実験条件として比較することで、それぞれが対話意欲の変動要因としてどのような影響を及ぼすか分析した。分析の結果、1台のロボットと対話するよりも2台のロボットと対話する方が対話意欲が高くなることが明らかとなった。また、ジェスチャ表出がないロボットと対話するよりもジェスチャ表出を行うロボットと対話する方が対話意欲が高くなることが明らかとなった。特に、対話ロボットの複数台化に伴うネガティブな影響について分析し、非言語情報を補足情報として補うことができることを示した。

今後の展望として、2台のロボットをどのように用いることで対話意欲が高まるシステムが構築できるかを検討する。

参考文献

- [1] 丸山洸太, 大和淳司, 杉山弘晃, ”複数台対話ロボットとの対話におけるユーザの対話意欲の変動要因分析,” HCG シンポジウム 2018, I-1-8, 2018.
- [2] 飯尾尊優, 吉川雄一郎, 石黒浩, ”ロボットの複数体化が対話感に及ぼす影響: 展示会におけるボタン入力対話体験の評価,” The 30th Annual Conference of the Japanese Society for Artificial Intelligence, 2016, 2016.
- [3] 杉山弘晃, 目黒豊美, 吉川雄一郎, 大和淳司, ”複数ロボット間連携による対話破綻回避効果の分析,” The 31st Annual Conference of the Japanese Society for Artificial Intelligence, 2017.
- [4] 前澤瑠星, 北條駿伸, 大和淳司, 杉山弘晃, ”複数の対話ロボット利用による対話破綻緩和効果の一人称視点での評価分析,” 2017 HCG シンポジウム, 2017.
- [5] 有本庸浩, 吉川雄一郎, 杉山弘晃, 目黒豊美, 大和淳司, 石黒浩, ”対話焦点の不連続性を隠蔽する複数ロボット間連携の生成,” 2016 人工知能学会研究会資料 SIG-SLUD-B505-13, pp.51-52, 2016.
- [6] 萩野谷浩美, 佐伯由香, ”ストレス指標における唾液 α アミラーゼ活用の有用性,” 日本看護技術学会誌, Vol. 10, No. 3, pp.19-28, 2012.
- [7] 北條駿伸, 前澤瑠星, 大和淳司, 杉山弘晃, ”対話破綻における複数メディアによる一人称視点での評価変動の分析,” 2017 HCG シンポジウム, I-1-8, 2017.
- [8] 中村卓也, 西村祥吾, 川波弘道, 神原誠之, 萩田紀博, ”ユーザとの信頼関係構築を目的としたロボット対話におけるペーシングの評価,” IEICE-CNR2018-17, Vol.118, No.184, pp.33-38, 2018.
- [9] 横山祥恵, 山本大介, 小林優佳, 土井美和子, ”高齢者向け対話インタフェース -雑談継続を目的とした話題提示・傾聴の切替式対話法,” 音声言語研究報告情報処理学会, Vol.2010, No.4, pp.1-6, 2010.