

演出を用いた実世界情報の顕在化の研究

Directed Human-Robot Interaction for Manifesting Real World Information

鳴海 真里子*¹ 今井 倫太*²
Mariko NARUMI Michita Imai

*¹慶應義塾大学大学院 理工学研究科

Graduate School of Science and Technology, Keio University

*²慶應義塾大学 理工学部, 科学技術振興事業団 さきがけ研究 21

Faculty of Science and Technology, Keio University

Precursory Research for Embryonix Science and Technology, Japan Science and Technology Corporation

This paper proposes “i-Director”, a dialog system on a humanoid robot “Robovie”, dealing with sensors’ information and actuator movement of the robot. i-Director generates an utterance which expresses person’s feeling or impression of a physical object or a situation in a real world and makes humans sympathize with robot. In this paper, we have conducted psychological experiment to confirm how “sympathy for robot” affects human-robot interaction using i-Director. The results of the experiment indicate that subjects sympathizing with the robot are significantly attracted to interaction with the robot.

1. はじめに

近年、多数の知能ロボットが開発されており、ロボットの持つコミュニケーション能力が注目されている。人間同士の対話では、情報伝達だけでなく「冷たいね」「気持ちいいね」といった感覚を表現した言葉が交わされている。例えば「あ、花が咲いている」「(その花は)きれいだね」という対話である。この対話は互いの感覚を察しながら行われ、人間同士の感覚の共有につながる。感覚の共有(以下「共感」と呼ぶ)は、コミュニケーションの上で重要な役割を果たす。ロボットがより高いコミュニケーション能力を得るためには、人間同士の対話と同様に人間が共感するコミュニケーションが必要である。本研究では、ジェスチャを含む音声対話による人間・ロボット間のコミュニケーションを扱う。

人間・ロボット間のコミュニケーションの研究は、人間・ロボット間の情報伝達を目標として、以前から行われている[1]。具体的には、人間がロボットに命令を与えたり、ロボットが人間に返答を返すといったことを実現している。

しかし、従来のコミュニケーションでは情報伝達を扱うことが可能であるのに対し、人間・ロボット間の共感についてはいまだに取り扱っていない。特に、人間のロボットに対する共感がコミュニケーションに与える影響については検証も行われていないのが現状である。人間同士が対話する場合、ジェスチャや視線によって構成された共同注意に基づいて無意識に相手の意図を推測し、共感が実現される。同様に人間がロボットと共感するには、ロボットが「人間との間に共同注意を構成し、人間にロボットの感覚を推論させる」という過程が必要となる[2]。

本稿では人間のロボットに対する共感がコミュニケーションに与える影響を調査するために、実験システム i-Director を開発した。i-Director は、実世界の状況や物に対して人間が共感しやすい発話(情報ではなく感覚を表現する発話)を生成できる。本研究では i-Director を用いて人間がロボットとのコミュニケーションに集中できるかを検証する。また上記の検証を基に、共感しやすい発話の影響によって人間がロボットの感覚を推測するかどうかに関して検証する。

2. 演出を用いた音声対話システム

本章では i-Director の概要について述べる。i-Director はコミュニケーションロボット Robovie[3] 上に実装された音声対話システムであり、「演出」を用いたインタラクションを実現する。演出とは、ロボットが注目している事象を人間に知らせるために行われるロボットの動作(発話及びジェスチャ生成)のことである。以下に、i-Director による発話及びジェスチャ生成について説明し、人間とロボットの共感の過程について述べる。

ロボットがある状況 S を感覚 F として認識したとき ($Robot(S, F)$) の人間が共感する過程について示す。i-Director は、演出(発話及びジェスチャ)により人間を自分と同じ状況 S に注目させたのち感覚 F を人間に伝える((1)式)。

$$(S, F) \quad direct(S), affect(F) \quad (1)$$

ここで、 $direct(S)$ は状況 S に注目させるための動作、 $affect(F)$ は感覚 F を表現する動作を表す。

(2)式に、(1)式の結果起こる共感について示す(S 及び F の添字 H と R は、それぞれ人間とロボットを表す)。(2)式では、 $direct(S)$ と $affect(F)$ によって人間とロボットの S と F が同一なものとなった結果、人間がロボットに共感していることを表している($Human(S, F)$)。

$$\begin{aligned} Robot(S_R, F_R), Human(S_H, F_H) \\ direct(S_R) \rightarrow S_R = S_H \\ affect(F_R) \rightarrow S_R = S_H, F_R = F_H \end{aligned} \quad (2)$$

例えば机の上に「花」があり「よい香り」であることを共感させるには、花を指して「花があるよ」と発話($direct(flower)$)し、「いい香りですよ」と発話($affect(good_smell)$)する。

2.1 共感の実現

i-Director を用いて前述の「共感」の定義に基づく動作をロボットが行った場合に想定される人間とロボットの状態を図1に示す。

人間は感覚器官を用いて実世界の物体を認識すると同時に、無意識のうちにその物体に対する感覚を得る。ここでロボット

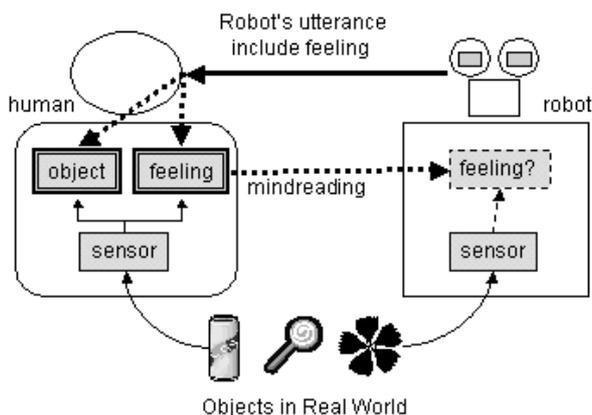


図 1: ロボットに対するマインドリーディング

が同じ物体に関する感覚表現 (例: 「おいしいね!」) を発話すると、人間はロボットに対してマインドリーディングを行って本来感覚を持たないはずのロボットの感覚を察する必要性が生じる。この他者の感覚を読むという行動によってインタラクションへのひきこみが起こる。

一方、図 1 の状況でロボットが物体に関する情報表現 (例: 「飴だよ」) を発話しても、人間はただ単に発話と実世界の物体を結びつけるだけで、ロボットへのマインドリーディングを行う必要性が生じない。結果、共感表現の発話に比べてインタラクションへのひきこみが起こりづらくなる。

i-Director による実験では、ロボットの発話の違い (感覚表現と情報表現) によって被験者のインタラクションへの集中の違いを検証する。つまり、被験者はインタラクション中も「これは実験である」と意識しているが、ひきこみによってその意識が薄れるとロボットとのインタラクションに集中する (図 2)。

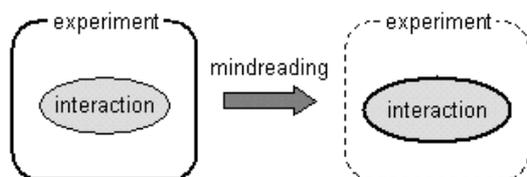


図 2: 実験における共感の影響

2.2 コミュニケーションロボット Robovie

i-Director が実装されたコミュニケーションロボット Robovie は、コミュニケーション機能に重点を置いて設計された擬人化が容易なヒューマノイド型ロボットである。Robovie の概観を図 3 に、ハードウェア構成を表 1 に示す。

2.3 i-Director の構成

i-Director は演出モジュールと対話モジュールで構成される。演出モジュールは与えられた (S, F) に基づいて選択された $direct(S) \cdot affect(F)$ の実行を管理する。対話モジュールは演出モジュールによって決定された行動を実行し、結果をフィードバックする。

対話モジュールは、演出モジュールから受け取った状況 S または感覚 F ・情報 I に基づいて行動 (発話及びアクチュエータの動作) を決定・実行し、与えられた時間制限 T 内で得られた



図 3: Robovie

表 1: Robovie の構成

| | |
|---------|--|
| アクチュエータ | 頭部 (3 自由度)、腕 (4 自由度) 2 輪独立駆動方式の車輪・キャスト |
| センサ | ステレオマイクروفोन 全方位視覚センサ ステレオカメラ (両眼の位置、 2 自由度で注視方向が制御可能) 接触センサ (上半身 18 個下半身 10 個) 超音波センサ (24 個) 赤外線センサ・温度センサ |

センサ入力 (音声認識・接触センサ) から最も重要な意味を持つものを選択・抽象化し、演出モジュールに返す。音声認識には Linux 2.2.16-rtl2.2 上で動作させた汎用大語彙連続音声認識エンジン Julian[4] を用いた。

3. 行動実験

本章では、人間のロボットに対する共感が持つ影響を検証するために i-Director を用いて行った行動実験について述べる。実験では、「人間はロボットとの対話によって「ロボットと共感した」と感じられる場合、ロボットとのコミュニケーションを作為的にとらず、コミュニケーション自体に集中できる」という仮説の検証を行う。

3.1 実験方法

本実験の具体的な手続きは以下のとおりである。

被験者:

男女の大学生 21 名の被験者を、実験群 11 名・対照群 10 名にランダムに配分する。

実験環境:

図 4 に実験環境を示す。本実験は、慶応義塾大学理工学部内の 4 階または 5 階にある教室において行われた*1。

実験条件:

表 2 に示すトピックについて、ロボットが「感覚的な発話」を行う条件 (実験条件) と、「情報伝達」を行う条件 (対照条件) を設定した。以降、各条件の被験者は実験群と対照群と呼ぶ。実験条件別のインタラクションは、図 4 の 1 の位置で、周辺の机の上に置かれた小道具 (図 4 の A~C) を用いて行われる。その後表 3 に示す各条件に共通なインタラクションが、図 4

*1 階数による環境の差異はない。

表 2: 実験条件別の演出トピック

| トピック | 感覚的発話例 | 情報発話例 | 小道具 (位置) | ロボットの位置 |
|-------|----------|---------|---------------|---------|
| 天気 | あったかいね | 明日も晴れるよ | - | 1 |
| 缶のお茶 | 冷たいでしょ | 烏龍茶だよ | お茶 (A) | 1 |
| お菓子 | おいしいでしょ | チョコだよ | お菓子 (ロボットの手先) | 1 |
| 花 | いい香りでしょ | バラだよ | バラの花 (B) | 1 |
| クッション | やわらかいでしょ | クッションだよ | クッション (C) | 1 |

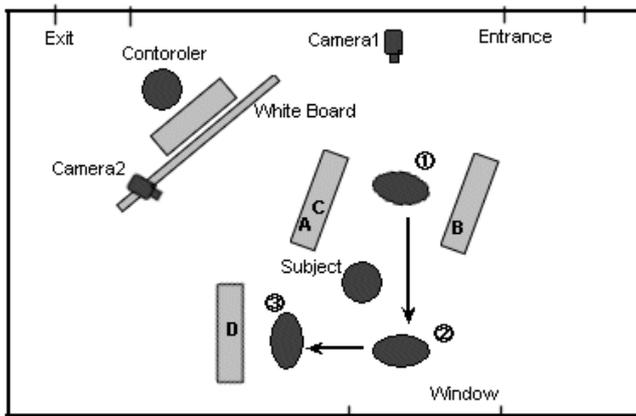


図 4: 実験環境

の 2 と 3 の位置に移動し、小道具 D を用いて行われる。ロボットの操作者はホワイトボードで隠されており、被験者からは見えない。

表 3: 各条件共通の演出トピック

| トピック | 発話例 | 小道具 (位置) | 位置 |
|------|-----------|-----------|----|
| 景色 | いい眺めでしょ | - | 2 |
| おみやげ | 好きな方を取ってね | 2 つの箱 (D) | 3 |
| 出口 | あっちが出口だよ | - | 3 |

実験手順:

実験は、以下の 4 ステップで行う。

1. 実験者は被験者をロボットの前に連れていき、「ロボットの相手をするように」と教示して退出する。
2. ロボットが動作を開始し、実験条件によって異なる対話シナリオに基づいて被験者とインタラクションを行う。
3. 各実験条件に共通な対話シナリオでインタラクションを行う。対話シナリオの最後で、ロボットが被験者に出口から退出するよう指示する。
4. 退出した被験者は、別室でロボットとのインタラクションについての質問票に答える。

評価方法:

評価は、被験者の行動観察と質問票への回答を用いて行う。質問票では、ロボットの発話の理解度・自分自身のロボットに対する反応とその理由について、被験者に回答を求めた。

3.2 仮説と予測

本実験では、以下の仮説に基づく 3 つの予測の検証を行う。

[仮説]

ロボットとの対話によって「ロボットと共感した」と感じられる場合、実験群の被験者はロボットとのコミュニケーションを作為的にとらえず、コミュニケーション自体に集中できる。

[予測 1]

いくつかの演出を行った後、ロボットがお菓子を差し出し「食べてみてね」と発話する演出を行う。実験群の被験者がそれまでのロボットの感覚表現の発話によって「ロボットと共感した」と感じられる場合、ロボットから渡されたお菓子を実際に食べる確率は、情報表現の発話を行った対照群の被験者と比較して増加する。

[予測 2]

各条件に依存したインタラクションを行った後、ロボットが窓に近付き「いい眺めでしょ」と発話する演出を行う。実験群の被験者がそれまでのロボットの感覚表現の発話によって「ロボットと共感した」と感じられる場合、ロボットの発話に同意する確率は、情報表現の発話を行った対照群の被験者と比較して増加する。

[予測 3]

実験群の被験者がそれまでのロボットの感覚表現の発話によって「ロボットと共感した」と感じられる場合、ロボットが渡した「おみやげ」を持って帰る確率は、情報表現の発話を行った対照群の被験者と比較して増加する。

3.3 実験結果

[共感の実現]

予測の検証に入る前に、仮説において前提となっている「ロボットと共感した」かどうかという点について検証する。(2) 式で示した共感とは、被験者がロボットの感覚に対する発話に同意することにより成立する。被験者の行動観察によると実験群の被験者全員がロボットの感覚的な発話に発話またはうなづきで同意していたことから、定義された共感とは成立していたといえる。

[予測 1]

行動観察によって得られた、お菓子を食べた被験者の数・食べなかった被験者の数を表 4 に示す。

この調査結果を分析したところ有意な差が確認された ($p < 0.05$) ことから、演出によりロボットに共感した実験群が、対照群よりもお菓子を食べることが確かめられた。

表 4: お菓子の演出に対する被験者の反応

| | 食べる | 食べない |
|-----|-----|------|
| 対照群 | 5 | 5 |
| 実験群 | 11 | 0 |

[予測 2]

行動観察によると、景色に関するロボットの発話の前後で実験群は 11 人全員、対照群は 1 人を除く 9 人が窓の外を眺めて景色を確認していた。景色を見た後、ロボットの発話に同意した被験者の数・同意しなかった被験者の数を表 5 に示す。同意しなかった被験者の多くは、ロボットの発話に答えず無言であった。

表 5: 景色に関する発話に対する被験者の反応

| | 同意する | 同意しない |
|-----|------|-------|
| 対照群 | 4 | 5 |
| 実験群 | 10 | 1 |

この調査結果を分析したところ有意な差が確認された ($p < 0.05$) ことから、演出によりロボットに共感した実験群が、対照群よりもロボットの景色に関する発話に同意することが確かめられた。

[予測 3]

行動観察によると、ロボットがおみやげについてジェスチャー及び発話で示した後、両群とも全員の被験者が 2 つのおみやげのうちどちらか 1 つを手にとった。さらにその後、ロボットからの“おみやげ”を持って部屋を退出した被験者の数・机に戻して持って帰らなかった被験者の数を表 6 に示す。

表 6: おみやげに対する被験者の反応

| | 持って帰る | 持って帰らない |
|-----|-------|---------|
| 対照群 | 2 | 8 |
| 実験群 | 7 | 4 |

この調査結果を分析したところ差が少ないながらも有意な差が確認された ($p < 0.1$) ことから、演出によりロボットに共感した実験群が、対照群よりもロボットからのおみやげを持って帰る傾向があるといえる。

質問票でおみやげを持って帰らなかった理由を尋ねたところ、「実験だから」「また後で使うと思ったから」というように実験であることを意識した回答が多く見られた。

以上の結果から、実験群の被験者は対照群の被験者よりも「ロボットの相手をする」という役割を確実に演じており、インタラクションへひきこまれていたといえる。

4. 考察

本章では、実験結果を踏まえ i-Director による共感が人間・ロボット間のコミュニケーションに与える影響について考察を行う。

「ロボットの景色に関する発話に同意する (予測 2)」という行動は、ロボットが感覚を持つことを人間が想定した結果であるといえる。また「ロボットからもらったお菓子を食べる (予測 1)」「ロボットからもらったおみやげを持って帰る (予測 3)」という行動は、ロボットの「(お菓子を) 食べてほしい」「(おみやげを) 持って帰ってほしい」という意図を人間がくんで行われたものである。

よって、i-Director によってロボットに共感した被験者は、無意識にロボットに対してマインドリーディングを行うことによって、ロボットが感覚を持つことを想定し、意図を積極的に推測したと考えられる。

さらに、おみやげを持って帰らなかった被験者が特に「実験である」ことを意識していたことから、図 2 に示したマインドリーディングによるインタラクションへのひきこみが起こっていたといえる。

演出の各トピックで被験者が発話する確率は実験群の方が高かった ($p < 0.05$)。ただし、これは対話シナリオの関係上実験群の方がロボットの発話に対する返答を求められる機会が多いことの影響も考えられる。対照群の被験者の発話確率は演出が進むほど下がっていったが、実験群の被験者の発話確率は実験の最後まで高い数値で推移していた。これはロボットの発話に対して常に何らかの返答をすることで対話が成立し、よりインタラクションにひきこまれていた結果といえる。

5. むすび

本稿では、人間とロボットの間で共感を成立させるといって観点から行動実験を行い、人間とロボットのインタラクションの分析を行った。また、共感を実現する実験システムとして、音声対話システム i-Director を実装した。実験結果より、人間はロボットと共感することによってロボットに対してマインドリーディングを行い、コミュニケーションへひきこまれることがわかった。

今後は演出トピックの拡張及び大規模な行動実験を行い、共感によるインタラクションへのひきこみについて検証する予定である。

参考文献

- [1] Toshihiro Matsui, Hideki Asoh, John Fry, Youichi Motomura, Futoshi Asano, Takio Kurita, Isao Hara and Nobuyuki Otsu: "Integrated Natural Spoken Dialogue System of Jijo-2 Mobile Robot for Office Services", Proceedings of AAI-1999, pp.621-627, 1999.
- [2] Tetsuo Ono, Michita Imai: "Reading Robot's Mind: A Model of Utterance Understanding based on the Theory of Mind Mechanism. Proceedings of AAI-2000, pp.142-148, 2000.
- [3] Takayuki Kanda, Hiroshi Ishiguro, Tetsuo Ono, Michita Imai and Ryohei Nakatsu: "Development and Evaluation of an Interactive Humanoid Robot "Robovie", IEEE International Conference on Robotics and Automation (ICRA 2002), pp.1848-1855, 2002.
- [4] 李晃伸, 鹿野清宏: "複数文法の同時認識および動的切り替えを行う認識エンジン Julius/Julian-3.3", 日本音響学会研究発表会講演論文集, 3-9-12, 秋季, 2002.