人間の五感を利用したロボットとのコミュニケーションへの没入の実現

今井 倫太*・鳴海 真里子**

Immersion in Interaction with a Robot via Human's Five Senses

Michita IMAI* and Mariko NARUMI**

This paper proposes a new communication strategy for engaging humans in interaction with a robot. In usual human-robot interaction, a person looks on the robot's behaviors without responding to them and seldom immerses him/herself in the interaction. Our communication strategy achieves person's immersion in interaction by generating an affective utterance which stimulates a person's feeling about a physical object. The innovative idea of the strategy is to employ human's five senses of the actual physical objects to mediate the human-robot interaction. In particular, since the affective utterance relates directly to a sort of human sensory perception, it is easy to develop a robot system to stimulates the person's feeling.

Key Words: humanoid-robot, human-robot interaction, immersion in interaction, interaction via physical world objects

1. はじめに

人間の生活環境で人と直接接するロボットが開発され、人間の日常生活をさまざまな側面で支援すると期待されている。人間の生活環境では一般の人とロボットがコミュニケーションすることとなり、特殊な知識を前提としていないユーザインタフェースが求められている。前提知識無しに使えるもっとも簡単なユーザインタフェースは言葉やジェスチャによるコミュニケーションシステムである。言葉やジェスチャは、母国語の違いを除けば誰もが利用可能なコミュニケーション手段だといえる。本稿では、音声やジェスチャを用いた人間とロボットのコミュニケーションについて取り扱う。

音声やジェスチャを交えたコミュニケーションでは、ロボットが話す内容や身振り手振りが人間の心理的な面にさまざまな形で影響する。受けた影響下で無意識の内にロボットに対する返答や行動を人間は変えて行くので、ロボット用のコミュニケーションシステムの設計では、人間に与える影響を考慮することが重要である。一方、ロボット用のコミュニケーションシステムの実現には、音声認識・合成や、自然言語解析、ジェスチャ認識、ロボットの行動生成といった要素技術を構築しなくてはならない。しかし以上の要素技術の完成度を上げるだけでは人間とロボットのコミュニケーションを成

立させることができない。人間とのコミュニケーションの最中にロボットが何を話してどのように振る舞うのかといったコンテンツが、人間の心理的側面に基づいてデザインされていることが重要である。たとえば、音声認識・合成の性能がたとえ優れていたとしても、ロボットの話す内容を人間が聞いてなければ対話は成立しない⁶⁾。さらに、周囲の状況についてロボットが説明する際にも人間の注意を説明対象へ適切に誘導できなかつたら、ロボットが何を話しているのか人間はわからない⁴⁾。以上の事態は、ロボットを展示会場に展示する際に頻繁に起こる。音声やジェスチャを交えた円滑なコミュニケーションを人間とロボット間で成立させるためには、人間への心理的な影響を検討しつつ、ロボットの発話やジェスチャをデザインし、ロボットとのコミュニケーションへと人間を引き込まなければならない。

人間への心理的な影響を考慮する形で人間とロボットのコミュニケーションを扱った研究はたくさんある。これらの研究の多くは、言葉やジェスチャまたは双方を用いた研究である。中でもロボットのジェスチャが人間に与える影響を考慮した研究^{2), 5), 7)}は最も基礎的な影響を利用している。これらの研究では、人間が、相手のジェスチャや表情、視線から興味の有無や注意対象を無意識の内に見出すことを利用している。ただし、設計されるジェスチャや表情、視線は人間に違和感をもたせない作りこみが必要である。たとえば、興味ある物体に対して視線を向け興味深そうに身を乗り出すジェスチャが考えられる。また、ロボットの発話とジェスチャ、視線を同時に使用して人間の注意を目的の対象へ誘導する研究も行なわれている^{4), 9)}。アイコンタクトや視線の動き、指さしジェスチャによって人間の注意を誘導しつつ、指示語を用

* 慶応義塾大学, 科学技術振興機構 さきがけタイプ

** 慶応義塾大学 (現所属: ソニーデジタルネットワークアプリケーションズ (株))

* Keio University and JST Presto

** Keio Univ. (Sony Digital Network Applications Inc.)

(Received August 22, 2005)

(Revised February 7, 2006)

いて人間に対象物の情報を伝える。また、人間とロボットの関係や人間とコンピュータの関係がコミュニケーションに与える影響を調べた研究もある^{3), 6)}。たとえば、小野等の研究では、人間とロボットの間に適切な関係が確立されていないと、人間は、ロボットの言葉を理解できないことを示した。具体的には、見ず知らずのロボットかどうか(親しい関係かどうか)で、人間がロボットの発話を理解できたりできなかったりする現象である。関係に依存して生じるこれらの現象は、人間が、相手のコミュニケーションの意図を推測する際に自分と相手の関係を参照していることを示している。

しかし興味の有無を示すジェスチャ・発話^{2), 5), 7)}や注意の誘導^{4), 9)}にはロボットの発話内容を人間が真剣に捉えるようにさせる効果がない。ここで、相手の話す内容を真剣に捉えることは、相手からの頼みごとを聞いたり、相手の勧めに従うことである。もしくは、聞かなかったり従わない場合には、きちんと断りを入れることである。人間同士のコミュニケーションでは、ロボットに代表される人工システムに比べ比較的容易に真剣なやりとりが起こる。しかし、われわれは、相手が人間であっても突然話しかけられたら、真剣に受け止めない場合が多々ある。そこで人間は、挨拶や世間話を交えたり、言葉やジェスチャの表現を工夫しながら、真剣な応対を相手から巧みに引き出そうとする。とはいっても、人間同士のコミュニケーションに比べると、ロボットとのコミュニケーションの難しさは、比較にならないほどの差がある。人間とロボットのコミュニケーションの場合には、人間がロボットに注目していたとしても、真剣なやりとりが起こり辛いのである。実は、ロボットに対して人間が注目する点が特異なために、真剣な受け答えの成立が難しくなっている。人間は、ロボットとインタラクションする際にロボットの動きや機能を観察することに集中しロボットの発話の内容を無視しがちなのである。小野等の研究⁶⁾の結果はこの現象を実験で再現したものである。ロボットの発話の内容を人間に理解させるためには、ロボットの観察を人間に止めさせ、ロボットの発話の意図を人間に推測させる必要があることを示唆している。小野等の研究では、ロボットがもつディスプレイにパーソナル CG エージェントを表示することによりロボットが自分と関係あると人間に思わせる。この構築された関係が意図を推測させるきっかけを作る。しかし、コミュニケーションロボットの一般的な形態である人型ロボットを考えると、パーソナル CG エージェントといった特殊なインタフェースではなく発話やジェスチャを用いたより一般的な手法が望ましいといえる。つまり、真剣なやりとりを人間とロボット間で実現するためには、人間に与える心理的な影響を考慮して、ロボットの意図を人間に推測させるための発話・ジェスチャをデザインする必要がある。

本稿では、ロボットの発話・ジェスチャを用いたコミュニケーション戦略を提案し、ロボットの発話を人間が真剣に捉えるコミュニケーションを実現する。本稿では、人間がロボットの発話を真剣に捉えている状態を、コミュニケーションへ

引き込まれた状態と呼ぶ。また、本コミュニケーション戦略によってロボットの意図を人間が推測ようになる課程をモデル化し、被験者実験によって検証する。

コミュニケーション戦略の中心的なアイデアは、実世界の物体や実世界の状態に対する人間の感覚をロボットの発話内容とすることである。ここで実世界の物体に対する感覚とは、人間が飲んだり・食べたり・触ったり・嗅いだり・見たりするときに五感で感じる、物体の種類に応じた感覚である。たとえば、人間がジュースを飲んで感じる冷たさがこの感覚にあたる。実世界の状態に対する感覚とは、気温や湿度に起因するものである。本コミュニケーション戦略は、人間の五感を利用するためにつぎの二つのステップから構成される。1) 発話・ジェスチャによって、実世界のある物体または実世界のある状態へ人間の注意を向ける、もしくは、実際に食べたり、飲んだり、触ったりするように促す。人間は、このステップで五感を通してある感覚をもつことになる。2) 前段のステップでの感覚を表現する発話をロボットが生成する。たとえば、人間がジュースを飲んでいる際にロボットが「冷たいでしょ」と発話する。ステップ 2) の発話は、対象の物体や状態に対して同様の感覚をロボットがもっていると人間に思わせることを狙っている。最終的に、人間は、ロボットの発話を解釈する際にロボットの感覚を推測すると思われるので、ロボットの意図も徐々に推測するようになる。

さらに、本コミュニケーション戦略は、ロボットの具体的な発話やジェスチャを容易にデザインできる。通常、言葉による高次なコミュニケーションでは、人間がさまざまな解釈をするので単純な手法でデザインできない。しかし、本コミュニケーション戦略は、人間の五感といったすべての人間に共通な情報を利用しているので、ステップ 1) で人間を注目させる状況に対してステップ 2) の発話を一意にデザインできる。人間を相手としたコミュニケーションで明確に発話内容を決定できることは通常希であることを考慮すると、本コミュニケーション戦略は、コミュニケーションへと人間を引き込む際の行動をデザインするにあたってたいへんすぐれた手法であるといえる。

本稿の構成は、つぎの通である。2章で人間とロボットのコミュニケーションの引き込みについて説明し、コミュニケーションの引き込みに必要となる要件についてまとめる。3章で、コミュニケーションの引き込みを実現するコミュニケーション戦略および戦略に基づいたモデルを提案する。4章で、被験者による検証実験について述べ、5章で実験結果について考察する。6章でまとめと今後の課題について述べる。

2. コミュニケーションへの引き込み

2.1 人間とロボットのコミュニケーション

人間とコミュニケーションロボットのコミュニケーションは、ロボットの発話を人間が真剣に捉えるかといった観点から分類すると大きく二つに分かれる (Fig. 1)。図では、ロボットが人間にお菓子をあげている場面である。図では、ロ

Fig. 1 Communication between a human and a robot

ロボットからの働きかけに対して真剣に「ありがとう」と人間が返答しているコミュニケーション（図右）と、何を返答していいのか分からずに無言になってしまうコミュニケーション（図左）の二通りが示されている。人間とロボットの理想的なコミュニケーションが成立しているのは、明らかに右側である。しかし、説明員がいない場所でロボットが人間とコミュニケーションする場合の多くは、左図に見られる理想的ではないコミュニケーションである。一方で人間同士では、無意識の内に理想的なコミュニケーションが確立される。

本稿では、ロボットとのコミュニケーションへと人間を引き込み、ロボットの発話を人間が真剣に捉えるようになる要件について考える。また、この要件を満たすロボットのコミュニケーション戦略を提案する。

2.2 心の理論と引き込み状態

Fig. 1 のコミュニケーションの違いは、人間が、ロボットとのインタラクションに没入しているかどうかの違いである。Fig. 2 は、Fig. 1 のコミュニケーションの違いをコミュニケーション中の人間の視点の違いで表わしている。Fig. 2 の左の四角は、ロボットとのコミュニケーションに人間が戸惑っている際の視点を表わしている。この視点では、ロボットの動きを人間が客観視 (calm viewpoint) しており、ロボットとのインタラクションであることを客観的に意識している。一方、Fig. 2 の右の四角は、人間がインタラクションに没入しているときの視点を表わしている。この視点では、インタラクションのやりとり自体に人間が集中しており、客観的視点は消えている。

つまり、理想的なコミュニケーションを実現するためには、ロボットとのインタラクションへの客観的な視点から、インタラクションに集中した視点へ人間の視点を変えることといえる。人間の視点を V と表わしまとめると、人間をロボットとのコミュニケーションへと引き込むということは、つぎのルールを実現することとなる。

[Rule1]

$V = \text{calm_viewpoint} \rightarrow V = \text{interaction_viewpoint}$

Rule 1 を実現するために、人間同士のコミュニケーション

Fig. 2 Difference of viewpoints of a human when he/she interacts with a robot

Fig. 3 Human-human and human-robot interaction

と人間とロボットのコミュニケーションをより詳しく比較する。Fig. 3 の (a) は、通常の間人同士のコミュニケーションを表わしている。人間は、他人の発話を理解する際に、相手の心を読み、発話の意図や、相手が知覚している状況を推測する^{1), 8)}。つまり、インタラクションに没入している時には、相手の発話の意図を推測するために相手の立場に立った視点を取っている。また、自分が返答をするために、相手の立場と同時に自分の立場の視点も取っている。この二の視点が同時に成立している状態がインタラクションに没入した視点である。インタラクションに没入した視点を取る結果、客観的な視点は消滅している。人間が誰かとコミュニケーションする際に相手の立場に立って相手の意図や知覚対象を推測することは、自閉症児の研究からも明らかになっており、心の理論と呼ばれる¹⁾。つまり、Fig. 3 の (a) は、心の理論を模式的に表わしたものと見える。

Fig. 3 の (b) は、人間が、ロボットとのインタラクションを客観視している場合のコミュニケーションを表わしている。人間がインタラクションを客観視している場合、ただ単にロボットが言葉を発話しても、人間は、ロボットの意図や知覚している対象を推測しないので、言葉だけを解釈し発話内容を理解することになる。しかし、言葉だけの解釈では、内容によって発話内容を理解できないことがある。さらに、ロボットとの関係性を人間に与えないと、人間は、ロボットの

Fig. 4 Communication robot Robovie

Table 1 Structure of Robovie

Actuators	Head (3 DOF), Arm (4 DOF) 2 drive wheel, one caster
Sensors	Stereo microphone Omni-direction vision sensor Stereo vision Touch sensors (Upper body: 18 pieces, lower body 10 pieces) 24 Ultra Sonic distance sensors Infrared sensors and thermometer

意図や知覚対象を推測せず、発話内容を理解できないことが実際に実験で確かめられている⁶⁾。

以上の人間とロボットおよび人間同士のコミュニケーションに関する考察を考慮すると、ロボットとのコミュニケーションへ人間を引き込むためには、ロボットの意図や知覚対象を人間に推測させるように、ロボットの立場の視点を人間にもたせるコミュニケーション戦略が必要となる。つまり、視点 V に対して以下の式が成り立つ必要がある。

$$V = \text{robot_viewpoint} \quad (1)$$

2.3 Robovie

本研究では、コミュニケーション戦略を実装するロボットとして Fig. 4 に示すコミュニケーションロボット Robovie を用いた。Robvie は、Table 1 に示すハードウェアで構成される。

Robovie の上半身は、人間に近い視線の動きジェスチャを生成できる自由度で構成される。人間とほぼ同じ形態をロボットがもち、同じ動きを生成できることは、ロボットの意図や知覚対象を人間が推測する際に人間のアナロジーを用いることができることを意味する。よって、人間と同じ意図や知覚対象をロボットにも想定しやすくなっている。この想定しやすさは、ロボットとのコミュニケーションへと人間を引き込む際にも有利に働く。

3. コミュニケーション戦略

ロボットとのコミュニケーションへと人間を引き込むためにロボットのコミュニケーション戦略を提案する。本コミュニケーション戦略は、つぎの二のステップで構成される。

ステップ 1. ある実世界の物や実世界の状態に対して人間

を注目させる発話やジェスチャを生成する。もしくは、物体を食べたり、飲んだり、触ったりするように人間を促す。

ステップ 2. 五感を通して人間がある感覚をステップ 1. で感じているので、この感覚に関係する発話をロボットが生成する。

本コミュニケーション戦略は、ステップ 1. で誘発する人間の感覚をステップ 2. の発話で刺激しているといえる。発話による刺激の結果、ロボットも同様の感覚をもっていると人間に思わせる。この過程を通して (1) 式で示したロボットの立場の視点へと人間の視点を導く。本稿では、ステップ 2. で生成される発話を感覚的発話 (affective utterance) と呼ぶ。

3.1 人間の五感の利用

本コミュニケーション戦略の二のステップを形式的に式で表わすと以下の通りとなる。

$$\text{Direction}(\text{Physical_event}), \text{Utterance}(H_feeling) \quad (2)$$

physical_event は、実世界のイベント (人間が物体に対して行なったこと) または周囲の状態を表わす。たとえば、現在の天候といった状態や、捕食といった人間の行動を挙げることができる。また、 $\text{Direction}(\text{Physical_event})$ は、 Physical_event への注目や、 Physical_event を起こすよう人間を促す発話を生成することを表わす。たとえば、天気への注意喚起や、食べるのを勧めるといったことである。 $H_feeling$ は、人間が感じている感覚であり、 $\text{Utterance}(H_feeling)$ は、 $H_feeling$ を表現する発話の生成を表わす。(2) 式は、 $\text{Direction}(\text{Physical_event})$ と $\text{Utterance}(H_feeling)$ がそれぞれステップ 1. と 2. に対応している。

本節では、(2) 式によるコミュニケーション戦略によって人間の視点に変化する過程を Fig. 5 (a) の形でモデル化する。Fig. 5 (a) は、 $\text{Direction}(\text{Physical_event})$ が終わり、実世界の物体や状態 (Physical_event) に対してある感覚を人間がもっている場面である。ここで、ロボットが、人間がもっているであろう感覚 $H_feeling$ に基づき感覚的発話 ($\text{Utterance}(H_feeling)$) を生成している。たとえば、ロボットが $\text{Direction}(\text{Physical_event})$ で人間にアメを食べさせたのなら、 $\text{Utterance}(H_feeling)$ は、「美味しいでしょ」といった発話となる。人間は、感覚的発話 $\text{Utterance}(H_feeling)$ を受け取ると、自分の感覚と結びつける (Fig. 5 (a) 人の絵の feeling への矢印)。自分の感覚と結びつけたつぎに、人間は、実世界の物体や状態に対して人間と同様の感覚能力をロボットももつと思うことになる (Fig. 5 (a) 人の絵の mind-reading の矢印)。Fig. 5 (a) のモデルでは、実世界の物体や状態に対してロボットが感覚をもつことを $\text{Eval}(\text{Physical_event})$ で表わす。

つまり、Fig. 5 (a) のモデルは、以下のルールに示す過程を人間に引き起こす効果がコミュニケーション戦略にあることを示している。

[Rule2]

$$\text{Utterance}(H_feeling) \rightarrow \text{Eval}(\text{Physical_event})$$

Fig. 5 Outline of directed interaction

さらに、ロボットと同じ感覚を共有していることを推測する過程で、人間の視点が $V = robot_viewpoint$ を取ることになる。つまり、以下のルールが成立する。

[Rule3]

$Eval(Physical_event) \rightarrow V = robot_viewpoint$

本コミュニケーション戦略が人間に与える効果の結果、人間の視点がロボットの立場にも置かれることになり、ロボットとのコミュニケーションへと人間は引き込まれる。

3.2 情報を表現した発話によるインタラクション

本節では、本コミュニケーション戦略および Fig. 5 (a) のモデルの効果を明確にするための、比較対象となる戦略およびモデルを説明する。この比較対象は、次章の被験者実験で本コミュニケーション戦略の比較対象として実際に用いる。

Fig. 5 (b) が対象モデルである。Fig. 5 (b) は、実世界の物体や状態に関する情報を表現した発話を感覚的発話の代わりに用いた場合のモデルである。本稿では、この発話を情報の発話と呼ぶ。つまり、Fig. 5 (b) は、ロボットが (2) 式の変わりに以下の対象戦略に従って発話を行なった場合を表わしている。

$Direction(Physical_event), Utterance(Object)$ (3)

ここで $Object$ は、 $Direction(Physical_event)$ によって人間が注目している実世界の物体や状態を表わす。 $Utterance(Object)$ は、実世界の物体や状態 ($Object$) に関する情報を伝える発話である。たとえば、Fig. 5 (a) の説明で用いたアメに関する例では、 $Utterance(Object)$ は、「アメだよ。」(商品名でもよい) といった情報提示となる。

Fig. 6 Experimental environment

対象戦略の場合には、Fig. 5 (b) のモデルが示す通り人間の視点がロボットの立場を取ることはない。なぜなら、人間は、 $Utterance(Object)$ による情報発話を受け取ると、自分が認識している $Object$ へと結びつけるからである (Fig. 5 (b) 人の絵の $Object$ への矢印)。 $Object$ への注目は、 $Direction(Physical_event)$ によりすでに確立されたものであるので、 $Utterance(Object)$ による情報的発話を解釈する際に人間の視点をロボットの立場に新たに置く必要がない。よって、視点の変更が起こりづらいつつ同時に、コミュニケーションへの引き込みも起こりづらい。

4. 被験者実験

本章では、本コミュニケーション戦略がもつ影響を検証するために、実験システムを構築し、被験者実験を行なう。2章で説明したとおり、コミュニケーションへの引き込みは、コミュニケーションの最中の人間の視点の違いとして表われる。そこで被験者実験では、以下の仮説の検証を行なう。

[仮説]

ロボットとの対話において感覚的発話を与えられた場合、人間はロボットとのコミュニケーションを客観的な視点でとらえず、コミュニケーション自体に引き込まれる。

4.1 実験環境

Fig. 6 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なわれた^(注1)。ロボットは、2章で紹介したロボット Robovie を用いた。Fig. 6 の番号 (1) から (3) は、被験者とコミュニケーションする際に Robovie が立つ場所を表わす。A (冷たい缶のお茶) と B (プレゼント) は、Robovie が $Direction(Physical_event)$ を実行する際に用いる小道具である。また、Robovie は、右手の先にチョコレートをあらかじめもっている。

Robovie 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。

4.2 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。

(注1) 階数による環境の差異はない。

Fig. 7 View from the window which is at the position (2) in Fig. 6

Table 2 Topics and directed utterances selected depending on the experimental group

Topics	Utterances based on Communication Strategy Step 1 for both groups (Behavior)	Affecting utterances for the experimental group which are based on Step 2.	Informative utterances for the control group which are not based on Step 2.	Referred equipments (its locations)	Robovie's location
Weather	Today	is warm.	is fine.	Non	(1)
Cold tea	A tea can is there. Please drink it. (Robovie Pointed to the tea can.)	It's cool.	It's SOUKENBI tea	Tea (A in Fig. 6)	(1)
Sweet	I'll give this. Please eat it. (Robovie handed a sweet.)	It's delicious.	It's KitKat.	Sweet (On Robovie's hand)	(1)

Table 3 Topics and directed utterances common to both groups

Topics	Utterances for both groups (Behavior)	Equipment (its location)	Robovie's location
Landscape	It's beautiful, isn't it? (Robovie directed its head to the window to see view from the window.)	Non	(2)
Present	This is a present for you. Take a favorite one. (Robovie pointed to one of the boxes)	Two boxes (B in Fig. 6)	(3)

4.3 実験条件

Table 2 に示すトピック (天気・お茶・お菓子) について, Robovie が「感覚的発話」を行なう条件 (実験条件) と, 「情報的発話」を行なう条件 (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。

4.4 発話シナリオ

実験において Robovie が発話する内容は, Table 2, 3 に示すシナリオの通りである。シナリオには, 五つの発話トピックがある。Table 2 に示した最初の三つのトピック (天気・お茶・お菓子) には, 二通りのシナリオが存在しており, シナリオに従って別々の発話が実験群と対照群に与えられる。一方が感覚的発話であり, 他方が情報的発話である。Table 3 に示した二つのトピック (景色・おみやげ) は, 感覚的発話と情報的発話の効果を検証するために用意したものであり, 実験群と対照群に共通して与えられる。つまり, 景色とおみやげのシナリオに対して被験者が真剣に対応するかどうかを検証するための物である。

最初の三つのトピックにおける感覚的発話は, 本論文で提案したコミュニケーション戦略に従ってデザインされたものである。Robovie は, ステップ 1 で, 世界の物に対して, ある感覚を被験者がもつよう仕向ける発話をする。引き続きステップ 2 で感覚的発話をする。たとえば, お茶のトピックでは, 「お茶があるよ。飲んでみてね」とお茶を飲むように勧め (ステップ 1), 続いて「冷たいね」と発話する (ステップ 2)。お菓子のトピックでは, 「これあげるね。食べてみてね」とお菓子を食べるように勧め (ステップ 1), 「美味しいよ」と感覚的な発話をする (ステップ 2)。ただし, 天気のトピックに関しては, 挨拶文の性格上ステップ 1 とステップ 2 が一つの発話文の中に含まれた形となっている。つまり天気のトピックは, 「今日は,」 (ステップ 1) と「温かいね」 (ステップ 2) となっている。

情報的発話の方は, コミュニケーション戦略のステップ 2 を, 情報を表現した発話に変更したものである。ステップ 1 の発話文は, 先の例と同一である。Table 2 で示す通り, 天

気のとピックに対して「天気いいね」、お茶のとピックに対して「壮健美茶だよ」、お菓子に対して「キットカットだよ」と情報を伝える発話をする。

後半の二つのとピック(景色・おみやげ)に関するシナリオは、Robovieの発話を被験者が真剣に受け止めるかを判定するためにデザインされたものであり、本研究で提案したコミュニケーション戦略に従ったものではない。具体的には、つぎの通りである。景色のとピックでは「良い眺めでしょ」という発話で、被験者に同意して欲しいというRobovieの意図を表わしている。この発話で、ロボットの意図通りに被験者が同意するかを検証する。ここで、被験者とRobovieが見る景色は、Fig. 7の写真に示す通り東京の都心を一望できるものであり、誰が見ても良い眺めという発話に同意できるものである。また、おみやげのとピックでは、Robovieが被験者におみやげを与える。Robovieから貰ったおみやげを、本当のおみやげとして被験者が真剣に受け取るかどうかという点を検証する。

4.5 実験手順

実験は、以下の4ステップで行なう。

- (1) 実験者は被験者をRobovie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) Robovieが動作を開始し、最初の三つの対話シナリオを実行し、被験者とインタラクションを行なう。対話シナリオは、実験条件に依存してTable 2の第3カラムと第4カラムのどちらか一方が選ばれる。このインタラクションは、Fig. 6の(1)の位置で、机上に置かれた小道具A (Fig. 6)を用いて行なわれる。また、手先のチョコレートも用いる。ここで、節4.4で説明した通り、お茶やお菓子を食すように勧めるのは、Robovieである。用いるのは、Table 2のステップ1の発話である。
- (3) その後Robovieは、Table 3に示す各条件に共通な二つのインタラクションを、Fig. 6の(2)と(3)の位置に移動し行なう。また、おみやげのとピックでは、小道具Bを用いる。対話シナリオの最後で、Robovie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Robovieとのインタラクションについての質問票に答える。

実験群に与えられる最初の三つの対話シナリオは、Table 2が示す通り、本研究で提案したコミュニケーション戦略でデザインされている。そして、つづくTable 3の発話は、その効果を調べるものである。したがって、この実験は、提案のコミュニケーション戦略に、ロボットとのコミュニケーションへと人間を引き込む効果があるかを調べるものである。

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なう。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。

Fig. 8 The subjects' response in the topic of sweet: EX Grp. denotes an experimental group and CTR Grp. denotes a control group

4.6 予測

ここで仮説に基づき実験結果を予測する。Table 2の最初の三つの実験条件別の対話シナリオにおいて、感覚的発話の効果が表われるのは、最初から三つ目の対話シナリオからであると思われる。そこで、Table 2の最後のとピック(お菓子)および、Table 3の二つの対話シナリオに対して被験者がとる行動の予測を述べる。

[予測1: お菓子のとピック]

お菓子のとピックでは、ロボットがお菓子を差し出し「食べてみてね」と発話する。実験群の被験者は、以前の対話シナリオでRobovieから感覚的発話を与えられているのでRobovieの発話を真剣に捉え、渡されたお菓子を実際に食べる確率が、情動的発話を与えられた対照群の被験者と比較して増加する。

[予測2: 景色のとピック]

景色のとピックでは、Robovieが窓に近付き「いい眺めでしょ」と発話する。実験群の被験者は、以前の対話シナリオでRobovieから感覚的発話を与えられているのでRobovieの発話を真剣に捉え、Robovieの発話に同意する確率が、情動的発話を与えられた対照群の被験者と比較して増加する。

[予測3: おみやげのとピック]

おみやげのとピックでは、Robovieがおみやげを指さし「おみやげがあるんだ、好きな方を取ってね。」と発話する。実験群の被験者は、以前の対話シナリオでRobovieから感覚的発話を与えられているのでRobovieの発話を真剣に捉え、Robovieから渡された「おみやげ」をもって部屋を退出する確率が、情動的発話を与えられた対照群の被験者と比較して増加する。

4.7 結果

[予測1に対する結果]

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Fig. 8に示す。また、実験シーンをFig. 9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、感覚的発話によりRobovieの発話内容を真剣に捉えるようになった実験群が、対照群より

Fig. 9 Experimental scenes of the sweet topic: the left picture shows an experimental group subject ate the sweet and the right shows an control group subject did not

Fig. 10 The subjects' response in the topic of landscape: EX Grp. denotes an experimental group and CTR Grp. denotes a control group

Fig. 11 Experimental scenes of the landscape topic: the left picture shows an experimental group subject agreed with Robovie's utterance and the right shows an control group subject did not respond to it

もお菓子を食することが確かめられた。

[予測 2 に対する結果]

行動観察によると、景色に関する Robovie 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。景色を見た後に Robovie の発話に同意した被験者の数・返答しなかった被験者の数を Fig. 10 に示す。また、実験シーンを Fig. 11 に示す。返答しなかった被験者の多くは、Robovie の発話に答えず無言であった。

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、感覚的発話により Robovie の発話内容を真剣に捉えるようになった実験群が、対照群よりも Robovie 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 に対する結果]

行動観察によると、Robovie がおみやげについてジェスチャーおよび発話で示した後、両群とも全員の被験者がおみやげを手にとった。さらにその後、Robovie からのおみやげをもっ

Fig. 12 The subjects' response in the topic of present: EX Grp. denotes an experimental group and CTR Grp. denotes a control group

Fig. 13 Experimental scenes of the present topic: the left picture shows an experimental group subject brought the present and the right shows an control group subject left it

て部屋を退出した被験者の数・机に置いてもって帰らなかった被験者の数を Fig. 12 に示す。また、実験シーンを Fig. 13 に示す。この調査結果を分析したところ有意傾向が確認された ($p < 0.1$) ことから、感覚的発話により Robovie の発話内容を真剣に捉えるようになった実験群が、対照群よりも Robovie からのおみやげをもって帰る傾向があるといえる。また、質問票でおみやげをもって帰らなかった被験者に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 Robovie の発話を真剣に捉えており、コミュニケーションへと引き込まれていたといえる。よって、仮説は示された。

5. 考 察

本章では、人間・ロボット間のコミュニケーションに本コミュニケーション戦略が与える影響について実験結果を踏まえ考察を行なう。「Robovie の景色に関する発話に同意する (予測 2)」という行動は、被験者が、Robovie の立場の視点で Robovie の感覚を推測していたと思われる結果である。また「Robovie からもらったお菓子を食する (予測 1)」「Robovie からもらったおみやげをもって帰る (予測 3)」という行動は、Robovie の「(お菓子を) 食べてほしい」「(おみやげを) もって帰ってほしい」という意図を被験者が汲んでいたと推測で

きる。また、意図の汲み取りおよび視点の変更と共に、被験者は、Robovieの発話を真剣に捉えるようになり、コミュニケーションへの引き込みが成立している。

また、アンケート結果では、おみやげをも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、客観的な視点でRobovieとのコミュニケーションを意識していたといえる。

さらに、各対話シナリオで被験者が発話する確率は実験群の方が高かった($p < 0.05$)。特に、対照群の被験者の発話確率は対話シナリオ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Robovieの発話を真剣に捉え、常に何らかの返答をしていたことを示す。つまり、よりRobovieとのインタラクションに引き込まれていた結果といえる。

6. ま と め

本稿では、ロボットとのコミュニケーションへと人間を引き込むためのコミュニケーション戦略を提案した。本コミュニケーション戦略は、二つのステップで構成される。ステップ1は、飲食物を飲んだり、食べたり、嗅いだり、触ったりするように人間を促すためのロボットの発話・ジェスチャである。もしくは、実世界の状態(気温、湿度)に注目させる発話・ジェスチャである。これらの発話・ジェスチャは、五感を通して人間の感覚を誘発することを狙っている。ステップ2で、誘発された感覚に関連する感覚的発話をロボットが生成する。人間は感覚的発話を与えられると、ロボットも人間と同一の感覚をもつと推測するようになると考えられ、推測を通して徐々にロボットの立場でロボットの意図や知覚対象も推測しはじめる。結果、人間は、ロボットの発話を真剣に捉えるようになり、コミュニケーションへの没入が実現される。

さらに、感覚的発話は、人間の感覚をダイレクトに表現したものであるため、ステップ1の内容に応じて簡単にデザインすることができる。つまり、個別のシナリオを容易にデザインできることを意味しており、コミュニケーションロボットのコンテンツ作成の基本的な設計論としても優れている。

また、本稿では、被験者実験によって本コミュニケーション戦略の効果を検証した。実験の結果、感覚的発話を与えられた被験者は、ロボットの発話を真剣に捉え、コミュニケーションへと没入することが確かめられた。

今回は、本コミュニケーション戦略を実験システムとして実現した。今後、実際のアプリケーションに組み込み、実環境での有効性についても検討する予定である。

参 考 文 献

- 1) S. Baron-Cohen: *Mindblindness: An essay on Autism and Theory of mind*. Mit Press (1987)
- 2) C. L. Breazeal: *Designing Sociable Robots*. MIT Press (2002)
- 3) B. リープス, C. ナス: 人はなぜコンピュータを人間として扱うか「メディアの等式」の心理学, 翔泳社 (2001)
- 4) M. Imai, T. Ono and H. Ishiguro: Physical relation and expression: Joint attention for human-robot interaction. In *IEEE Transactions on Industrial Electronics (ITIED 6)*, 50, 636/643 (2003)
- 5) H. Kozima, C. Nakagawa and Y. Yasuda: Interactive robots for communication-care: A case-study in autism therapy. In *Proc. of 14th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2005)*, 341/346 (2005)
- 6) T. Ono and M. Imai: Reading a robot's mind: A model of utterance understanding based on the theory of mind mechanism. In *Proceedings of AAAI-2000*, 142/148 (2000)
- 7) B. Scassellati: Quantitative metrics of social response for autism diagnosis. In *Proc. of 14th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2005)*, 585/590 (2005)
- 8) D. Sperber and D. Wilson: *Relevance: Communication and Cognition*. Oxford: Basil Blackwell (1986)
- 9) O. Sugiyama, T. Kanda, M. Imai, H. Ishiguro and N. Hagita: Three-layered draw-attention model for humanoid robots with gestures and verbal cues. In *Proc. of 2005 IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS 2005)*, 2140/2145 (2005)

[著 者 紹 介]

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。94年同大大学院計算機科学専攻修士課程修了。同年、日本電信電話(株)入社。97年ATR知能映像通信研究所へ出向。2002年慶應義塾大学大学院開放環境科学専攻博士課程修了。現在、慶應義塾大学理工学部情報工学科 助教授および、科学技術振興機構さきがけタイプ「相互作用と賢さ」研究員、ATR知能ロボティクス研究所客員研究員。ロボットとの対話、センサを用いた状況知覚に興味をもつ。IEEE、ACM、情報処理学会、電子情報通信学会、人工知能学会、ヒューマンインタフェース学会、認知科学会各会員。博士(工学)。

鳴 海 真 里 子

2003年慶應義塾大学理工学部情報工学科卒業。2005年同大大学院開放環境科学専攻修士課程修了。同年、ソニーデジタルネットワークアプリケーションズ(株)入社。第65回情報処理学会全国大会大会優秀賞受賞。ヒューマン・ロボットインタラクションに興味をもつ。情報処理学会会員。