

状況に引き出された発話による対話の形成とその心理的評価

鈴木紀子[†] 竹内勇剛[†]
石井和夫^{†*} 岡田美智男[†]

日常的な対話では、大局的な課題や目標の達成を目的とした課題指向的な側面だけではなく、むしろ対話の過程や話題の展開そのものを愉しむという側面がある。対話者間では、状況に応じた発話を積み重ねていくことにより、対人的な関係を形成し、それを維持していく。そうした過程を通して、対話者間で共通の理解を創り上げる。日常的な対話では、課題や目標の達成に先立ち、局所的で状況に依存した関係の形成とその調整が重要であると考えられる。このような視点から筆者らは、対話者間での局所的な関係の形成と維持を第一義的なものとしてとらえ、それをマルチエージェントの相互作用からなる計算機構によって実現するインタラクティブシステム「Talking Eye」の構築を進めている。本論文では、この Talking Eye システムと被験者の間での関わりを通じた対話者間での局所的な関係形成とその維持について、心理的指標に基づいた評定実験を行った。Talking Eye システムは、大局的な課題遂行のための対話管理メカニズムをまだ備えていない。そのため、本評定実験においても、被験者による課題遂行についての心理的評定は低いという結果を得た。しかしながら、対話過程に対する心理的評定は必ずしも低くはないことを確認した。さらに、局所的な関係形成を考慮した Wizard of Oz 法に基づく評定実験を行った結果、同様の傾向が確認された。これらの結果から、人同士の場合と同様に、人と Talking Eye システムの間においても局所的で状況に依存した関係が形成され、維持されることを明らかにした。

Organizing Situated Utterances of Dialogue and Its Psychological Evaluation

NORIKO SUZUKI,[†] YUGO TAKEUCHI,[†] KAZUO ISHII^{†,*}
and MICHIO OKADA[†]

Our everyday dialogue has not only a task-specific aspect that aims at the achievement of global tasks and goals, but also a social aspect that relates to enjoyment of the development of topics and the dialogue process itself. The relationship among participants is organized by the situated utterances that are derived successively from dialogue. Inter-subjectivity among participants emerges during this process in dialogue. We consider that the process of local relationship among participants is more important than the achievement of tasks on everyday dialogue. From this viewpoint, we have focused on the organization of local relationship among participants as the primary function of dialogue. Therefore, we have constructed an interactive system called "Talking Eye" that has an architecture in which utterances and motions are derived dynamically based on the local interaction among multi-agents. This paper presents a psychological experiment to evaluate the social behaviors of the Talking Eye system using the mental evaluations of subjects. Subjects evaluated the achievement of tasks negatively, because the Talking Eye system did not have the architecture for the dialogue management that could achieve global tasks. However, subjects evaluated the process of dialogue positively. Moreover, the same tendency was observed in interactions using the Wizard of Oz method. These results suggested that the local relationship organized and encouraged between subjects and the Talking Eye system was similar to that between humans.

1. はじめに

日々の生活の中で、我々は必ずしも明確な課題や目標を抱きながら行為を遂行しているわけではない。むしろ、周囲の状況に引き出されるように行為が出現し、そうした行為の系列の中で、本来の動機や目的があらわにされる場合が多い。このことは日常的な対話にも

[†] ATR 知能映像通信研究所
ATR Media Integration & Communications Research
Laboratories

^{*} 現在、ソニー株式会社
Presently with Sony Co.

あてはまる。日々の何気ない対話では、あらかじめ与えられた課題の達成を目的とした課題指向的な側面だけではなく、雑談などに代表されるように局所的な対話のやりとりや話題の展開そのものを愉しむという側面がある¹⁾。また対話者間の関わりの中で状況に引き出された発話が積み重ねられ、そうした過程を通して、対人的な関係が形成され、維持されていく。同時に、こうした関わりの中でコミュニケーションの本来の目的であった、対話者間での共通の理解が創り上げられる²⁾。この視点からいえば、日常的な対話場面での協調的な相互行為による問題解決は、局所的な関わりの結果として達成されるものといえる。その意味ではまた、日常的な対話では課題遂行や情報伝達の側面に先立ち、対話者間での局所的で状況に依存した関係の形成とその調整が重要であると考えられる。こうした側面は、インタラクティブシステムの設計などにおいても不可欠な要素であると考えられる。

筆者らは、あらかじめ与えられた課題の遂行を第一義的にとらえるインタラクティブシステムではなく、むしろ状況に引き出された発話を積み重ねることにより、対人的な関係を形成し、それを維持することを主眼としたインタラクティブシステム「Talking Eye」の構築を進めている^{3)~5)}。ここでは、マルチエージェントの相互作用の結果として、個々の発話や動作などの行為を状況の中から引き出すメカニズム⁶⁾を実現している。現段階での Talking Eye システムは、大局的な課題遂行のための対話の進行を管理するメカニズムをまだ備えていない。しかし、発話間での局所的な制約を満足させながら、状況に依存した形で個々の発話を引き出すために、対話者との関わりを通して対人的な関係の形成とその維持とを実現することができる。

本論文では、対話者との関わりを通して、状況の中で引き出された個々の発話や動作の振舞いにおけるシステムの親和的な側面を調べるために、心理的な指標を用いて評定実験を行い、その結果について考察を行う。この実験では、システムと被験者との関わりあいにおける、(A) 課題遂行、および、(B) 対話過程についての心理的な評定と、さらに、被験者との局所的な関係形成を考慮した Wizard of Oz 法⁷⁾に基づく心理的な評定によって、被験者と Talking Eye システムとの対話での評定値との比較検討を行う。

以下、2章では状況によって引き出される発話の積み重ねによる対話の形成過程について説明し、3章では、そうしたメカニズムを備える Talking Eye システムの概要を述べる。また、4章では被験者による Talking Eye システムとの関わりに対する心理的な評定の結果

とそれに対する考察を行う。5章では、4章で得られた知見を基に、今後の研究の展開について述べる。6章では本論文のまとめを示す。

2. 日常的な対話の形成

2.1 状況によって引き出される発話

日常的な対話には、相手に明確なメッセージを伝達することに主眼を置いたものがある。たとえば、事務手続きや商談では、対話者間で同じ課題を共有していることを前提に、その課題を遂行するための対話を行う。しかし日常的な対話では、このような課題遂行や目標達成のために明確なメッセージを相互にやりとりするものだけに限らない。むしろ、対話者間で共通した大局的な課題や目標を持たない状態で、互いに関わることがある。たとえば、友人との世間話や主婦の井戸端会議では、対話を行うことによる情報交換に加えて、対話者間の対人的な関係の形成・維持を行えること自体に意味があるといえる。このような対話では、その時々対話の状況や対話者間での局所的な相互作用の中で発話や動作が次々と引き出され、その積み重ねによって結果的に対話が形成・維持されていく¹⁾。

表1に示すように、これまで構築されてきた多くのインタラクティブシステムは、日常での商談や事務手続きの場合と同様に、人とコンピュータの間である課題を共有し、その課題を遂行することを前提に設計されている(本論文では、このような対話を「課題指向対話」とよぶ)^{8)~10)}。これらのシステムは、大局的な対話管理のためのメカニズムを備えており、課題遂行の効率性や操作性を向上させるための手段として音声言語が使われる。

一方、対話者との関係の形成・維持を実現することを目的として設計されたシステムに「ねね」¹¹⁾がある。このシステムは、感情モデルや対話理解の機構を用いて人との相互作用の過程で感情的な発話や動作を表出するメカニズムを組み込むことで、人と交感的な対話を実現しようとしたシステムである(このシステムに対する入力テキストによって行われる)。

筆者らは、対話の過程で対話者間での局所的な関係

表1 人-人、人-コンピュータ間の対話形態
Table 1 Interactive style between humans/human and computer.

	課題指向対話	交感的な対話
人-人	事務手続き 商談	世間話 井戸端会議
人-コンピュータ	情報提供システム ⁸⁾ 窓口業務システム ⁹⁾ 作業支援システム ¹⁰⁾	「ねね」 ¹¹⁾ Talking Eye ^{3)~5)}

が形成・維持されることにより、結果として人と交感的な対話を創り出すことを目指したインタラクティブシステム「Talking Eye」の構築を進めている^{3)~5)}。现阶段での Talking Eye システムは、大局的な課題遂行のための対話管理の機能をまだ備えていない。しかし、その時々対話の流れや相手の発話や動作に対して相互依存的な振舞いを産出するメカニズムを備えているため、その積み重ねによって結果として対話を形成・維持していくことが可能なシステムである。

2.2 発話の積み重ねによる対話の形成

課題指向対話における個々の発話は、目標達成のために一意に解釈が可能なメッセージとして、相手に正確に伝えられることを前提とする。そこでの対話はあらかじめ用意された枠組みに沿って進行する形式的なものだと見なされ、また対話者が話題に対して関わりあう視点は相互に知っており、対話者のいずれの発話に対してもその後の話題の展開の方向性は予測できる。そのため、人とコンピュータの間で課題指向的な対話を行うインタラクティブシステムは、予測される話題の展開の道筋に従って、適切な形で発話や動作を生成していく。一方、交感的な対話においては、そこで展開される話題は予測することができない。このような対話は参加者にそれぞれ帰属した動機に基づいており、各々の対話者が話題に対して関わりあう視点は異なっている。一般に、対話者が相互に関わる話題に対して、各々の発話が話題の展開にどのように寄与したかは次の発話がなされるまで知ることはできない。すなわち、ある時点における発話 U_n が、話題 T_n に対してどのような影響を与えるかについてはあらかじめ知ることはできない。発話 U_n の話題 T_n に対する影響が明らかになるのは、次の発話 U_{n+1} によって話題 T_{n+1} が変化したときである。つまり、発話 U_{n+1} によって話題 T_n を T_{n+1} に変化した U_n の影響は、発話 U_{n+1} によって初めて知ることができる。同様に発話 U_{n+1} は別の発話 U_{n+2} によって、対話者が相互に関わりあう状態を知りうる。このように、対話者は発話にともなう話題の展開のダイナミクスの中で発話し、相互に関わりあう話題は対話者間の局所的な相互作用の過程において立ち現れるといえる(図1)。

3. インタラクティブシステム Talking Eye

3.1 システムの構成

人とコンピュータとの間での局所的な相互作用の積み重ねの結果として対話が形成・維持されていくことを実証するために、Talking Eye システムの構築を進めてきた³⁾。このシステムは、3次元コンピュータグ

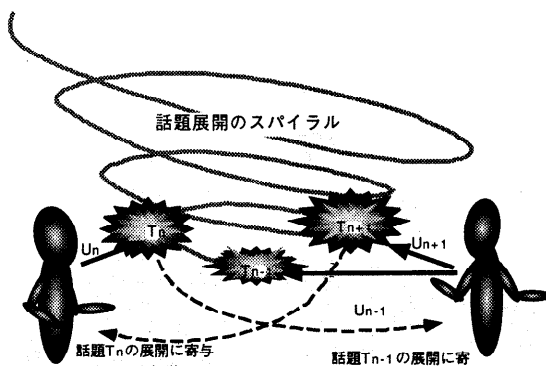


図1 局所的な相互作用による対話の形成

Fig. 1 Organization of dialogue through local interaction.

ラフィックスで描かれた「一つの目玉」というシンプルな外観を持った複数の仮想生物が仮想空間上で世界を創り、仮想生物どうして会話を自発的に展開していく。同時に外界の人と音声を通じて関わりあう。この仮想生物は、静止画で見る限り動物とも植物とも判断できない。しかし、ひとたび動き出すと多様な解釈が可能な振舞いを表出する(図2)。仮想生物がこのような抽象的な外観を持つことによって、特別な先入観を与えずに対話を成立させることができ、対話の過程に対して心理的な評価を行ううえで有効なデバイスとなっている。

Talking Eye システムの構成例を図3に示す(動的な行動選択モデルに基づく計算機構の記号の詳細については、図4参照)。本システムは、プロソディ検出(prosody detection)、動的な行動選択モデルに基づく計算機構(architecture based on dynamical action selection network)、CG生成(motion generation)、音声生成(phrase generation)のために1台、フレーズ認識(phrase recognition)に1台の計2台のワークステーションで構成されている。プロソディの検出では、音声の相対的なパワーのレベルや基本周波数の時間的変化を検出する。フレーズ認識では、単語グラフを用いた連続音声認識システム ATR SPREC¹²⁾を援用し、その認識語彙数は155単語である。言語制約は、FSA(Finite State Automaton)で記述されており、文節内文法で受理可能な発話の断片の任意の接続を可能にする。

3.2 動的な行動選択

対話者間での局所的な相互作用の中から発話や動作を導き出す機能を実現するために、筆者らは Agent Network Architecture (ANA) に基づいた動的な行動選択モデル⁶⁾を援用している(図4)。この動的な行動

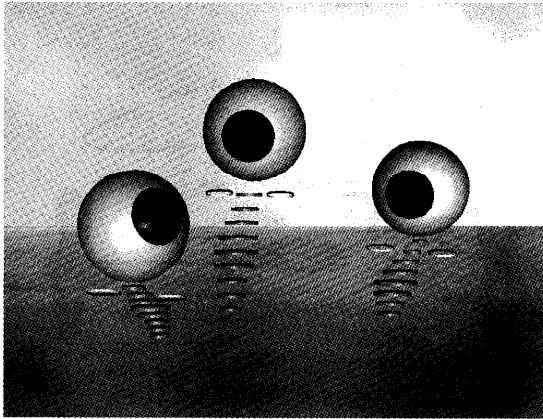


図2 「Talking Eye」システム

Fig. 2 Appearance of interactive system "Talking Eye."

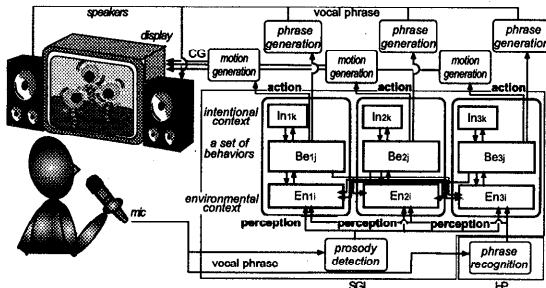


図3 「Talking Eye」システムの構成

Fig. 3 Configuration of the Talking Eye system.

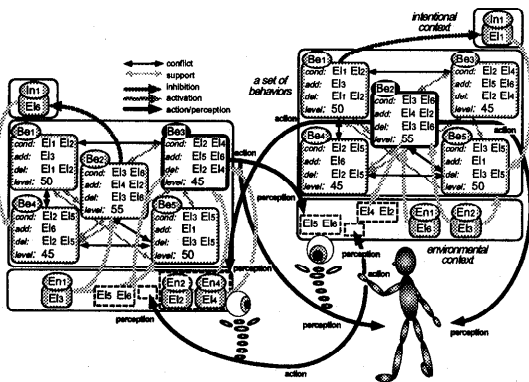


図4 外界との局所的な相互作用

Fig. 4 Local interaction with real world.

選択のためのモデルは、多数の「ふるまい (behavior)」とその環境との相互作用の中から、つねに変化する状況に適した「ふるまい」が動的に選択されるという特徴を備えており、人と Talking Eye システムの間で局所的な関係の形成・維持を実現するためのモデルとして適している。動的な行動選択のメカニズムを社会的な相互行為の場面に展開するために、個々の「ふる

まい」に対して新たな設定と記述を行った。特に、ロボットと環境との関係に見られるような「自律主体と環境との相互作用」ではなく、社会的な相互行為の場面では「同一の自律主体どうしの相互作用」となる。また、(a) 潜在的な多重のゴールが関わりの中から分節される、(b) 相手からの発話や動作などのイベントは対話の過程で事後的に分節される、(c) 相互行為の結果として対話が形成されるなど、従来の目的指向対話のモデルにはない、日常的な対話の性質を表現することに適したモデルとして実現される。

本論文ではシステムを取り巻く環境として、環境的文脈 (En_h : environmental context)、意図的文脈 (In_j : intentional context) とよぶ2種類の文脈をシステム内部に用意している³⁾。環境的文脈は、それまでの対話の状況や外界からのイベントを要素 El_k として持つ。意図的文脈は仮想生物の心的状態や対話に関する動機を要素 El_l として持つ。「ふるまい」の集合 (a set of behavior) を構成する個々の「ふるまい (Be_i)」には、外界に対して表出する行動が記述されている。

「ふるまい」には、基本的に⁶⁾における記法にのっとり、自身が外界に対して活性化するための局所的な制約が記述されている。図4に示されるように、「ふるまい」は、前提条件 (cond)、追加条件 (add)、削除条件 (del)、活性値 (level) のリストから成り立つ。前提条件は、「ふるまい」が活性化するために環境的文脈中に揃わなければならない条件が記述されている。追加条件 (add) には、「ふるまい」が活性化後に環境的文脈へ追加する条件が、削除条件 (del) には、「ふるまい」が活性化後に環境的文脈から削除する条件が各々記述されている。活性値 (level) は、その「ふるまい」がネットワークを通して伝搬させる活性化エネルギーの量、および「ふるまい」が活性化するために必要な活性化エネルギーの閾値を表す。さらに、「ふるまい」自体には活性化された際に外界に対して表出する発話や動作の種類が記述されている。

「ふるまい」どうしの間には、各々の局所的な制約に基づいた協調/競合関係があり、ネットワークを介した活性エネルギーの授受により局所的な相互作用を行う。すなわち、ある「ふるまい」 Be_1 にとって協調関係にある「ふるまい」 Be_2 に対しては、 Be_1 から Be_2 に対して活性エネルギーを与えることによって、 Be_2 を支援し、 Be_1 にとって競合関係にある「ふるまい」 Be_3 に対しては、 Be_1 から Be_3 の活性エネルギーを奪うことによって、 Be_3 を抑制する。また、外界との相互作用も同様に、環境的文脈の中に Be_1 の前

提条件と一致する要素が、あるいは意図的文脈の中に Be_1 の追加条件と一致する要素が存在する場合、各文脈から Be_1 に対して活性化エネルギーが与えられ、「ふるまい」 Be_1 の活性化が支援される。各「ふるまい」は、外界と「ふるまい」の集合および「ふるまい」どうしの相互作用の結果、自身の前提条件がすべて満たされ、かつ集積した活性化エネルギーの値がある閾値を超えた場合に活性化され、外界に対してその「ふるまい」に帰属する具体的な発話や行為が表出される。活性化後、その「ふるまい」が持つ追加条件は環境的文脈へ追加、削除条件は環境的文脈から削除され、さらに活性化エネルギーの値は初期化される。ただし、複数の「ふるまい」が集積した活性化エネルギーの値が閾値を越えた場合、その中で最大の活性化エネルギーを集積した「ふるまい」が活性化される。閾値を上回る活性化エネルギーを集積した「ふるまい」が存在しない場合は、活性化が起こらない。

本システムでは、400 (msec) を超える無音区間に挟まれた一続きの音声の連続を発話断片と定義し、イベントの検出を行っている。環境的文脈のイベントの要素として、プロソディ検出を用いた (i) 発話断片の相対的なパワーのレベル、(ii) 発話断片の末尾 500 (msec) の基本周波数の傾きの 2 種類と、フレーズ認識の結果を用いた (iii) 発話断片の意味的な素性および (iv) 他の仮想生物の発話の意味的な素性の 4 種類を用いている。各々のイベントの内容がそのまま環境的文脈の要素 $E1$ となる。プロソディ検出の結果は、発話断片の相対的なパワーのレベルの強/弱、基本周波数の傾きの上/下/平坦がイベントの内容として付与される。フレーズ認識の結果は、たとえば、「なにかしよう」「あそぼう」などは提案、「しずかにして」「こっちむいて」などが依頼など、意味的な素性がイベントの内容として付与される。

また、仮想生物の外界に対する行為としては、音声と動作の 2 種類があり、それぞれ相異なる「ふるまい」の集合を持つ。そのため、音声と動作の「ふるまい」は、動的な行動選択の原理を基に各々独立に選択される。仮想生物の音声として、同意、提案、否定、躊躇、妥協など約 50 種類の「ふるまい」が設定されている。仮想生物は外界との局所的な相互作用の中で、「そやなあ」「そんないやや」「こんなんどう?」「まあええか」「ほんま?」といった、あらかじめ録音された約 400 フレーズの簡単な日本語 (関西弁) の音声を発することができる。また、仮想生物の動作としては、大きくなる、首を横に振る、うなずきなどの約 20 種類が設定されている。これらは CG によって表現され、人は

これらの動きを見ることによって仮想生物の、驚き、失望、うなずきなどの心的状態を知覚できる。

Talking Eye システム上の仮想生物の音声や動作は、これらの機能を備えることで、人の音声言語に対してリアルタイムに誘発される。このような発話や動作の積み重ねの結果として対話が形成・維持されていく。

4. 心理的評価実験

4.1 背景

本章では、Talking Eye システムの被験者との関わりあいにおける親和的な側面を調べるために、心理的な指標を用いて評価実験を行い、その結果について考察を行う。

インタラクティブシステムを利用するうえでの心理的な指標に基づく評価は、これまでも行われてきた。その多くはシステムの特長上、主に課題や目標を達成するうえでの利便性や効率性、有用性、心理的負荷の多少などを評価するものであった^{13),14)}。この場合の評価は、ユーザがある課題を達成することを目的として行われるため、一般に課題の達成に支障があれば、意図やメッセージの伝達や解釈の過程に問題があったと見なされる。ところが対話の過程における交感的な側面に関しては考慮されていない。交感的な対話では、正確で効率的なメッセージの伝達は第一義的な問題ではない。むしろ関わりあうことによって対話者間で対人関係を形成・維持していく親和的な側面が重要であると考えられる。ここでは、人と Talking Eye システムとの関わりあいの過程における親和的な側面を心理的な指標を用いて評価することを試みる。

4.2 実験環境の設定

被験者は、21 インチのディスプレイに表示された 3 個体の仮想生物と音声を介して関わりあう (図 5)。

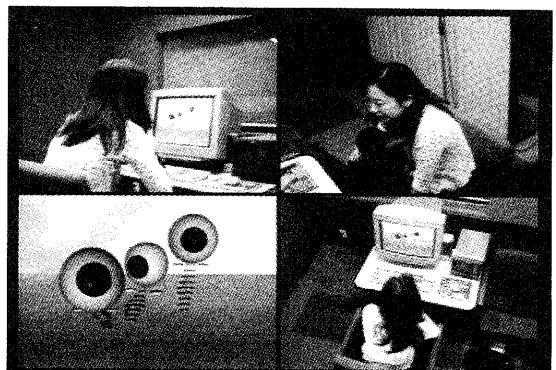


図 5 Talking Eye との対話の様子

Fig. 5 Example of human-Talking Eye system dialogue session.

3 個体の仮想生物との間に関わりあう実験環境を設定しているのは、以下の理由による。

- 1 個体と人の場合、半ば強制的に関わりあわなくてはならないような心理的負荷を被験者に与えてしまう。
- 2 個体と人の場合、2 個体の仮想生物どうしの関わりあいも可能なため、人がその輪の中に入るきっかけをつかむのが難しい。
- 3 個体と人の場合、仮想生物どうしが関わりあっても、そのうちの 1 個体はその中心的な参与者になっていない場合があり、その 1 個体が人と関わりあう対象となる機会を得る。

3 個体の仮想生物は各々異なった色（黄、青、桃）をしており、その行為は独立に動的な行動選択のモデルに基づいて制御される。被験者はマイクを通して音声によって仮想生物と関わりあう。一方、仮想生物は音声および CG 画像によって動作を表出する。ここでは、3 個体の持つ「ふるまい」の数、種類およびその記述内容はまったく同じものを用いている。しかし、環境的文脈に含まれているイベントの種類と各仮想生物の「ふるまい」は 1 対 1 対応ではなく、ある時点の仮想生物の内部状態（各「ふるまい」が持つ活性化エネルギーの値の推移の過程）によって、環境的文脈に含まれるイベントを分節する方法が変化する。仮想生物の内部状態は、外界の人および他の仮想生物どうしとの相互作用の結果として時々刻々と変化していくため、ある時点で 3 個体が同じ内部状態にある可能性はきわめて低くなる。そのため、人が話しかけた場合に、3 個体とも同じ発話の「ふるまい」が活性化される可能性も同様に、きわめて低いといえる。また、計算機の処理能力により、外界からの働きかけに対して各個体の環境的文脈が変化するタイミングが微妙に異なるため、3 個体が実世界に向けて同時に発話と動作を返すことはない。

4.3 教示および評価項目

被験者は、被験者自身の指示のとおりに関わりあうことが振る舞うように求めることを課題として与えられた。実際には小学校の教室などでの場面を想定し、被験者は教師で仮想生物は小学校低学年の子供たちという設定で、勝手に私語をしてやまない子供たちを静かにさせるという課題を教師役の被験者に与えた。ただし、この課題は仮想生物との間では共有されていない。すなわち、各仮想生物の意図的文脈には、課題達成のために協調的に振る舞うという局所的なゴールは設定されていない。逆に反抗的に振る舞うように意図的文脈を設定したため、仮想生物たちは反抗的な「ふるまい」

が活性化しやすい状態となった。例として、被験者が仮想生物に対して何か提案しても受け入れない、依頼しても拒否するなどがあげられる。被験者はこのような子供たちの注意を引き、言うことを聞かそうと様々な方法で関わりあいを試みる。しかし、結果的には課題で設定された目標は達成できない。このような実験環境の下で、以下の心理的な評定値を調べた。

(A) 課題遂行についての評価 被験者の課題遂行に関するシステムについての心理的評定。

【評価項目】

- 意図伝達度：どれだけ自分の意図が通じたか
- 目標達成度：どれだけ目標を達成することができたか
- 課題遂行上の協調度：どれだけ課題を遂行するうえで協調的だったか
- 課題遂行上の対応度：どれだけ課題を遂行するうえで対応が良かったか

(B) 対話過程についての評価 被験者の関わりあいの過程に対するシステムについての心理的評定。

【評価項目】

- 積極性：関わりあううえでどれだけ積極的だったか
- 親近性：関わりあううえでどれだけ親しみやすかったか
- 外向性：関わりあううえでどれだけ外向的だったか
- 友好性：関わりあううえでどれだけ友好的だったか

これら各 4 評価項目の評価値の平均が、それぞれの実験による観測対象の評価値とする。なお、すべての評価項目は 10 点尺度で求められる。評価値が大きいほど評価項目の内容を満たしている。これらの評価は Talking Eye システムとの関わりあいの終了後に調査用紙によって回答を得る。

4.4 実験

4.4.1 実験 1

実験 1 では、被験者と動的な行動選択のモデルに基づいて発話や動作が誘発される仮想生物との関わりあいを観測する。被験者は画面上に表示された 3 個体の仮想生物と音声を通して関わりあう。実験環境は 4.2 節に示したとおりである。

被験者 男女の大学生および大学院生 26 名

手順 各被験者に対し、以下の手順にて実験を施行する。

- (1) 事前に VTR に収録した人と仮想生物との関わりあいの様子を 1 分程度見せる。

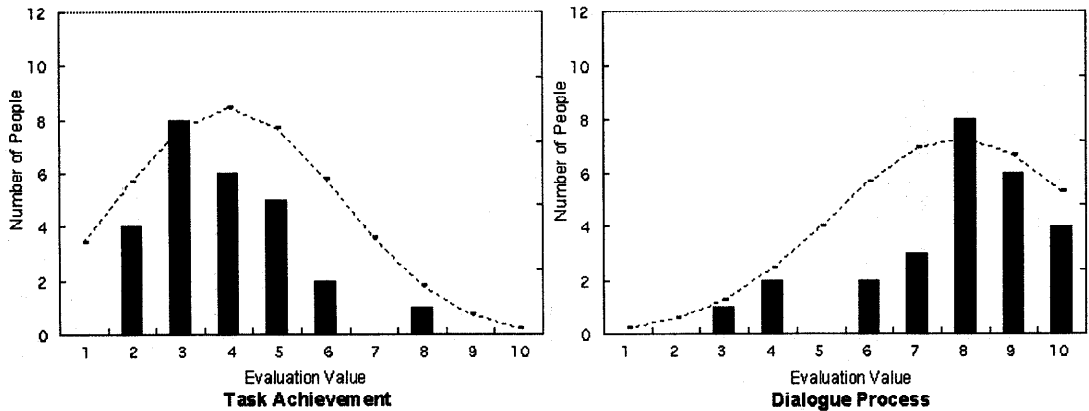


図6 実験1の結果：(A) 課題遂行について(左)，(B) 対話過程について(右)

Fig. 6 Histogram of evaluation value for experiment 1: (A) for task achievement (left), (B) for the dialogue process (right).

表2 実験1の人と Talking Eye システムの間における関わりあいに関する心理的評価

Table 2 Result of evaluation of interaction between human and Talking Eye system for experiment 1.

課題遂行		対話過程	
平均	SD	平均	SD
3.52	2.26	7.38	2.61

- (2) 実際に仮想生物との関わりあいの練習試行を1分程度させる。この間に実験に関する教示を行う。
- (3) 教示終了後、実験者は室外に退出し課題開始の合図(チャイム)を鳴らす。
- (4) 2分間の課題遂行時間後、任意に終了の合図を鳴らす。活発に関わりあっている場合には終了の合図を鳴らさず、しばらくそのままその進行に任せる。
- (5) 対話終了後、回答用紙に各評価項目に関する質問の回答を行う。

【実験結果と考察】

表2に人と Talking Eye システムの間での関わりあいに関する心理的評価の結果を示す。さらに、図6に横軸を評価値、縦軸をその評価値を与えた人数とした場合の心理的評価結果のヒストグラムを示す。この図は、評価値が5以下は負の方向の評価であり、6以上が正の方向の評価を示す。これらから、(A) 課題遂行についての評価が低くても (B) 対話過程についての評価は必ずしも低くならない、ということが明らかにされた。音声言語を目標達成のための情報伝達的手段として利用しているシステムでは、一般に (A) の評価の低さは (B) に強く影響すると考えられている。つまり目標の達成に障害があったとき、それに従属して

課題遂行の過程に障害の原因が帰属され評価が下がるのである。ところが実験1では、それぞれの評価は対照的な結果になった。このことは、関わりあうことが目標達成の手段として用いられたのではないことを示している。むしろ、人と仮想生物の関わりあいによって、互いの発話や動作が引き出されていき、その積み重ねの結果として親和的な関係が形成されたことを示唆している。付録表4に、被験者と Talking Eye システムとの対話例を示す。下線部で示したように、仮想生物の拒絶的な態度に対して、被験者の一般的な社会的反応とは異なる振舞いが観察された。これより、課題を達成するための関わりあいよりも、むしろ交感的な関わりあいを維持することによって、人と仮想生物の間で親和的な関係が形成されることが示されたといえる。

4.4.2 実験2

実験1によって、人と Talking Eye システムとの間で、局所的な相互作用の過程で引き出された発話や動作を積み重ねることによって、事後的にシステムの対人的に親和的な側面が見い出されたことを示唆する結果が得られた。しかしこの結果からだけでは、Talking Eye システム内の仮想生物の発話や動作が、局所的な相互作用の結果として引き出されていたためだという理由の根拠としては乏しい。そこで実験2では、Wizard of Oz 法 (WOZ 法) を用いて人と仮想生物の発話と動作を背後で操作する人とのインタラクションを観測した。この WOZ 法を用いることで、実験1で観測された対話が人と Talking Eye システムの間で固有なものだったのではなく、人同士の間で日常的に行われている関わりあいが実験1において同様に成り立っていたことを明らかにすることができる。

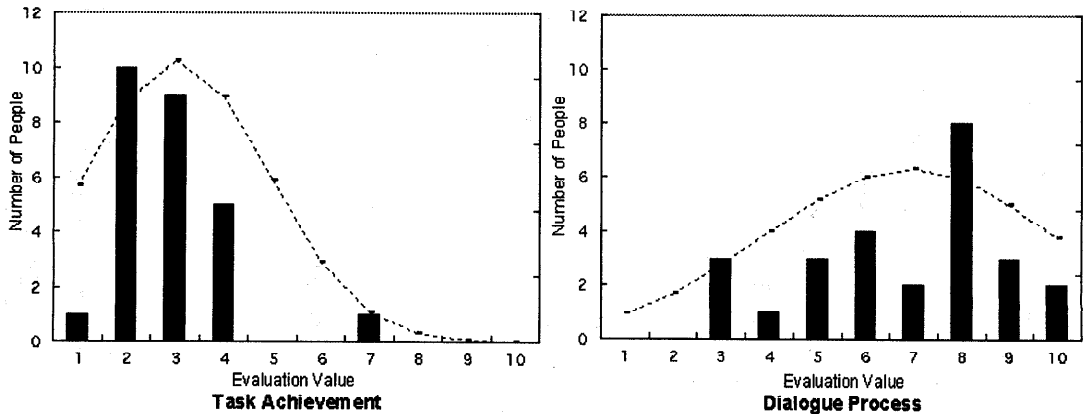


図7 実験2の結果：(A) 課題遂行について(左)、(B) 対話過程について(右)

Fig. 7 Histogram of evaluation value for experiment (2): (A) for task achievement (left), (b) for the dialogue process (right).

表3 実験2のWOZ法による人とTalking Eyeシステムの間における関わりあいに関する心理的評定

Table 3 Result of evaluation about interaction between human and Talking Eye system based on the WOZ method for experiment 2.

課題遂行		対話過程	
平均	SD	平均	SD
2.52	1.87	6.42	3.04

ここで用いるWOZ法とは、3個体の仮想生物の発話と動作を実験者のリアルタイムな操作によって選択することで、被験者がどのような反応をするかを観測する実験方法である。すなわち被験者は実験者と、仮想生物のために用意された音声と動作を用いて関わりあうことになる。本論文では、1名の操作者が対話の状況に応じて3個体の仮想生物の発話と動作を選択した。操作者は、各個体に対して独立な個性を仮定して発話や動作を選択するように指示されており、そのため、各個体に異なる操作者がいる場合と同等な、一貫性のある態度が産出されることが期待できる。

実験2に参加した被験者は実験1と同じである。ただし、実際には実験1の後に実験2を行った13名の被験者群と、実験2の後に実験1を行った13名の被験者群が存在する。このことにより、Talking Eyeシステムとの関わりあいの経験による学習効果を相殺できる。実験2の手順は実験1とまったく同様に行われた。

【実験結果と考察】

実験後の心理評定では、各個体間の態度についての調査を行った。その結果、実験1および実験2において、各個体間の評定値には明らかな差が観察されなかった。したがって、本実験で用いたWOZ法では、1

名の実験者がその振舞いを操作した場合でも、各個体はいずれも均一な態度を表出していたと考えられる。

表3に人とTalking Eyeシステムの間での関わりあいに関する心理的評定の結果を示す。さらに、図7に横軸を評価値、縦軸をその評価値を与えた人数とした場合の心理的評定結果のヒストグラムを示す。この図は、評定値が5以下は負の方向の評定であり、6以上が正の方向の評定を示す。この結果は、実験1と同様の評定傾向を示している。すなわち、(A) 課題遂行についての評定が低くても (B) 対話過程についての評定は必ずしも低くならない、という結果を得た。つまり、被験者は実験1においてTalking Eyeシステムと固有な対話を行っていたのではなく、実質的には人同士の場合と同様に局所的な相互作用によって引き出された発話や動作の積み重ねによって対話が形成・維持されていたといえる。

また、10点尺度を用いて子細な心理的評定を行ったため、実験1と実験2において、(A) 課題遂行についての評定および (B) 対話過程についての評定のそれぞれの平均と標準偏差 (SD) は、以下のような関係であることが分かる。

- (A) では、実験1の方が実験2に比べて平均値は大きいですがSD値は小さい。
- (B) では、実験1の方が実験2に比べて平均値は小さいがSD値は大きい。

それぞれの評価における平均値とSD値の違いは統計的に有意ではないが、傾向として次のような傾向が観測された。WOZ法を用いて仮想生物のふりをした人との対話は、システムに制御された仮想生物に比べて1ポイント近く低い評定を与えられている。これは、WOZ法の方が被験者の発話を正確に認識でき、さら

に的確な応答が可能だったため、被験者と仮想生物の操作者との間で課題指向対話が断片的に成り立っていたためと考えられる。すなわち、対話を手段として用いたときに、その目標を達成できなかった場合には、目標を達成する過程の関わり合いが課題遂行にうまく機能しなかったと解釈されやすい。

5. 今後の展開

本研究は、日常的な対話の特徴を構成的に明らかにすることによって、従来の課題指向対話や目的指向対話とは異なる日常的な対話のモデルの構築を試みている。このような視点から、発話や動作が状況に引き出されるメカニズムを備えたインタラクティブシステム Talking Eye を構築した。4章に示した心理的評価実験の結果が示すように、Talking Eye システムには、課題の遂行という実務的な能力が欠けているにもかかわらず、状況に応じた発話や動作を導出するメカニズムを備えていれば、関わり合いの過程で人はシステムに対して親和的な側面を見出しうることが分かった。このことは、コンピュータが社会的な存在として受け入れられるために、これまでの音声対話システムでは主に課題遂行における効率性や利便性を重視してきたが、同様に交感的な対話を実現するメカニズムもまた総合的な評価を高めるうえで重要な心理的要因となることを示唆している。

このような知見に基づく工学的な応用の可能性を以下にあげる。

- (1) 対話の過程において、人が対人的な志向性を見出しうる社会的実在の構築
- (2) 対話の結果として、人の心理的負荷を軽減する社会的実在の構築

(1)については、たとえば、交感的なインタラクションを実現するメカニズムを備えた社会的実在と、人の分身であるアバタが仮想社会に混在し、関わり合いを介して対人的な関係を形成していく場面への応用が考えられる。(2)については、関わり合いを通して社会的実在が乳幼児やペットのように人の心を癒す存在となる可能性が考えられる。さらに、これらの社会的実在を用いて、関わり合いの過程における対人的な志向姿勢や疑似対人行動の誘発の条件などの認知科学的な側面を探求することも可能となる。

6. むすび

日常的な対話においては、大局的な課題の遂行に先立ち、対話者間での局所的で状況に依存した関係の形成とその調整が重要であると考えられる。Talking

Eye システムは、マルチエージェントの相互作用からなる計算機構によって、状況を反映した発話を引き出す。また、この状況に引き出された発話を積み重ねることにより、対人的な関係を形成し、それを維持する。本論文では、被験者と Talking Eye システムの対話を通して、局所的な相互作用の中で導き出された発話や動作の振舞いにおけるシステムの親和的な側面を調べることを試みた。心理的な指標を用いた評価実験により、以下の結果を得た。

- (1) インタラクティブシステム Talking Eye と人との関わり合いにおいては、(A) 課題遂行についての心理的評価が低いという結果にもかかわらず、(B) 対話過程についての心理的評価は必ずしも低くならない。
- (2) 同様のことは Wizard of Oz 法を用いた場合においても観測され、人は Talking Eye システムとの間で、人同士の場合と同様の親和的な対話が行われていた。

これらの結果から、システムが大局的な課題遂行のための対話管理のメカニズムを備えていない場合でも、局所的な相互作用に基づいて引き出される発話や動作を積み重ねることによって、人同士の場合と同様に、人とシステムの間においても親和的な関係が形成され、維持される可能性があることを明らかにした。

今後は、本研究で得られた成果とあわせて、人とシステムとの関わり合いにおいて、共感をともなう関係生成のメカニズムを追求したいと考えている。

謝辞 本研究を行うにあたり、様々な面から援助していただいている ATR 知能映像通信研究所社長中津良平氏、有益な議論および励ましをいただいている同研究所第四研究室室長片桐恭弘氏をはじめ第四研究室の皆様にご感謝いたします。また、システムの実装に尽力していただいた古屋隆志氏、西垣茂生氏の両氏に深く感謝いたします。

参考文献

- 1) 岡田美智男：口ごもるコンピュータ，共立出版(1995)。
- 2) Schegloff, E.A.: Repair after next turn: The last structurally provided defence of intersubjectivity in conversation, *American Journal of Sociology*, Vol.97, No.5, pp.1295-1345 (1992)。
- 3) 岡田美智男：Talking Eyes: 対話する「身体」を創る，システム/制御/情報，Vol.41, No.8, pp.323-328 (1997)。
- 4) Suzuki, N., Inokuchi, S., Ishii, K. and Okada, M.: Chatting with Interactive Agent, *Eu-*

表4 実験1の対話例(25歳男性)

Table 4 Example of dialogue for experiment 1: 25 years-old, male.

H	おーい みんなこっち向いてや	
T1	なんやなんや	正面 → 上下に伸縮
T2	あっちいってーな	正面 → 上下に伸縮
H	まあ そんなこといわんと	
T3	うるさい	正面 → うなづく
T1	まあ ええやないか	正面 → 首を振る
H	ごめんごめん (pause) ちよつとだけだまって	
T2	なんでやねん	正面 → 首を振る
T1	まあ ええか	正面 → 首を振る
H	うん	
T1	くやしかったらここまできてみい	正面 → 首を振る
H	それはでけへんけどな (pause) ま 聞いてや	
T3	ほつといてーな	正面 → 上を向く
T2	いやや	正面 → 首を振る
H	まあまあ	
T3	なんやねんもう	正面 → 上を向く
T1	なんやと	正面 → 首を振る
H	まあまあ (pause) あそば	
T1	そやなあ	正面 → 上下に伸縮
H	あそばましょ	
T2	えーっ	正面 → 首を振る
T3	うーん	正面 → 上を向く

* pause : 500 (msec) 以上の無音区間

roSpeech '97, Vol.4, pp.2243-2246 (1997).

- 5) Suzuki, N., Ishii, K. and Okada, M.: Talking Eyes: Autonomous creature as accomplice for human, *APCHI '98*, pp.409-414 (1998).
- 6) Maes, P.: How to do the right thing, *Connection Science*, Vol.1, No.3, pp.291-323 (1989).
- 7) Malsby, D., Greenberg, S. and Mander, R.: Prototyping an intelligent agent through wizard of oz, *INTERCHI '93*, pp.277-284 (1993).
- 8) 中川聖一, 山本誠治: 音声対話システムの構成法とユーザ発話の関係, *信学論 (D-II)*, Vol.J79-D-II, No.12, pp.2139-2145 (1996).
- 9) Zue, V., Seneff, S., Polifroni, J., Phillips, M., Pao, C., Goddeau, D., Glass, J. and Brill, E.: PEGASUS: A spoken language interface for on-line air travel planning, *ARPA Workshop on Human Language Technology*, pp.196-201 (1994).
- 10) 安藤ハル, 北原義典, 畑岡信夫: インテリアデザイン支援システムを対象としたマルチモーダルインタフェースの評価, *信学論 (D-II)*, Vol.J77-D-II, No.8, pp.1465-1474 (1994).
- 11) 戸田正直: 心を持った機械—ソフトウェアとしての「感情」システム, *ダイヤモンド社* (1987).
- 12) Shimizu, T., Yamamoto, H., Masataki, H., Matsunaga, S. and Sagisaka, Y.: Spontaneous dialogue speech recognition using cross-word

context constrained word graphs, *ICASSP '96*, pp.145-148 (1996).

- 13) 伊藤敏彦, 小暮 悟, 中川聖一: 協調的応答を備えた音声対話システムとその評価, *情報処理学会論文誌*, Vol.39, No.5, pp.1248-1257 (1998).
- 14) Hugunin, J. and Zue, V.: On the design of effective speech-based interfaces for desktop applications, *EuroSpeech '97*, Vol.3, pp.1335-1338 (1997).

付録 被験者と「Talking Eye」システムの対話例

人(H)とTalking Eyeシステムの3個体の仮想生物(T1~T3)の対話例を表4に示す。

(平成10年10月9日受付)

(平成11年2月8日採録)

鈴木 紀子 (正会員)



1992年横浜国立大学工学部電子情報工学科卒業。1994年横浜国立大学工学研究科電子情報工学専攻修士課程修了。同年、ATR通信システム研究所入所。現在、ATR知能映像通信研究所研究員。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。電子情報通信学会、日本音響学会、人工知能学会各会員。

竹内 勇剛 (正会員)



1992年宇都宮大学工学部情報工学科卒業。1999年名古屋大学大学院人間情報学研究科社会情報学専攻博士後期課程修了。博士(学術)。現在、ATR知能映像通信研究所客員研究員。社会性に基づく人-モノ間のコミュニケーションに関する研究に従事。モノや人以外の生き物に対する「人らしさ」の帰属に関する問題に興味がある。認知科学会、人工知能学会各会員。



石井 和夫（正会員）

1984年神戸大学工学部電気工学科卒業。1986年同大学院工学研究科電子工学専攻修士課程修了。同年ソニー（株）入社。ビデオプリンタの開発、音声認識システムの研究開発に従事。1997年1月ATR知能映像通信研究所に出向。音声対話における対話の楽しさとコミュニケーションスキルの発達について研究。1999年1月ソニー（株）に復帰。人工知能学会、日本音響学会、電子情報通信学会各会員。



岡田美智男（正会員）

1982年宇都宮大学工学部電子工学科卒業。1987年東北大学大学院工学研究科博士課程修了。同年、NTT基礎研究所情報科学研究部。1995年よりATR知能映像通信研究所、主任研究員。工学博士。日常でのコミュニケーション場面における「身体」の主體的な振舞いとその環境、他者、メディアとの関わりに興味を持つ。著書に「口ごもるコンピュータ」（共立出版）等。認知科学会、人工知能学会、ソフトウェア科学会等会員。