

非分節音による反響的な模倣とその心理的影響

鈴木 紀子[†] 竹内 勇剛[†]
石井 和夫[†] 岡田 美智男[†]

人間は自分の発話や動作の一部を相手に真似されると、その相手の模倣的な振舞いの背後に自分に対して積極的に働きかける姿勢や感情を見いだしてしまう。その結果、相手に対する共感的な感情が生まれ、情緒的な関係の形成・維持が促進される可能性が高い。このことを検証するために本論文では、直前に発せられた発話を非分節音に変換し、音声の大きさ、リズム、イントネーションといったプロソディのレベルで模倣する機能を備えた仮想的なクリーチャをコンピュータ上に構築した。また、このクリーチャに対して発した音声を模倣されることによって引き起こされる心理的な影響について調べた。心理的評価実験では、直前に発した発話のプロソディを反響的に模倣する模倣的音声の占める割合と、一定のプロソディの音声を発する非模倣的音声の占める割合とが異なる3種類のクリーチャを用意し、各クリーチャとの相互作用に対する印象の評価を行った。その結果、被験者の発した音声に対する応答のうち、模倣的音声の占める割合の高いクリーチャほど、被験者のそのクリーチャの態度に対する好感度の評価値が高くなることを確認した。この結果は、自分の発話を反響的に模倣されることによって、非生物である仮想的なクリーチャに対しても共感的な感情をいだく可能性があることを示唆している。

Psychological Effects Derived from Echoic Mimicry Using Inarticulate Sounds

NORIKO SUZUKI,[†] YUGO TAKEUCHI,[†] KAZUO ISHII,[†]
and MICHIO OKADA[†]

In this paper we describe results from an experiment of an interaction with an artificial creature that mimics human voice echoically using inarticulate sound. We consider that humans are apt to find a partner's intention or emotion to themselves, when the partner mimics their utterance echoically under prosodic level. As a result, we regard that empathic interaction emerge among them. We test this hypothesis by having subjects interact with three artificial creatures that give different rates of their respective response, mimicked prosody voice or constant prosody voice. Subjects' evaluations of the creature were collected with a questionnaire according to impression of interaction with the creature. The result suggests that the higher the mimicry rate is for the creature, the more positive the impression of the subject is. We consider that the result supported our hypothesis that the echoic mimicry is a key for the emergence of empathic interaction between humans and computers.

1. はじめに

人間は自分の行動の一部を相手に模倣されると、その相手が自分に対して何らかの意図や感情を持っていると認知する傾向がある。たとえ九官鳥やおうむのように必ずしも人間に対する意図や感情を持つとは限らない相手であっても、自分の発した音声を文字どおり

「おうむ返し」されると、その背後に何らかの意図や感情を感じることがある。また、日常的な会話においても、相手に自分の発話の一部をなぞられたり、相手の発話を無意図的になぞって「おうむ返し」することがある。このような相手の何気ない動きに対しても、何か合理的な判断の下でなされた有意義な行動として理解しようとする。このような傾向は志向的な姿勢と呼ばれている¹⁾。人間は、志向的な姿勢の下で相手と関わる過程で、相手の模倣的な振舞いによって積極的な感情を引き出されることもあれば、消極的な感情を引き出されることもある。したがって、志向的な姿勢に基づいて相手の行動を意図的なものとしてとらえてし

[†] 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 Corporation

まう傾向は、相手との情緒的な関係を形成・維持する手がかりとなると考えられる。

筆者らは、コンピュータやインタフェースエージェントとの間において情緒的な関係を形成・維持するための方法を模索している²⁾。その手がかりの1つとして、自分の表出した振舞いを機械に模倣されることによってもたらされる心理的な影響に着目している。それを検証する手段の1つとして、これまで筆者らは互いに相手の発する音声を模倣し合うことで、相手に対する共感が引き出されていくことを目指したインタラクティブシステムを構築してきた^{3),4)}。このシステムは、外界から入力された音声の大小に関する時系列パターンを獲得し、特徴量に基づいて分節化していく。さらに、その結果として外界に向けて非分節音を用いた模倣的な音声を出力するメカニズムを備えている。

本論文では、コンピュータ上に構築した仮想的なクリーチャとの模倣的な音声を介した相互作用に着目し、クリーチャが外界から入力された音声を大きさ、リズム、イントネーションといったプロソディのレベルで反響的に模倣することによって生じる心理的効果を心理実験によって検討する。被験者が自分の発話を仮想的なクリーチャに反響的に模倣されることによって、クリーチャに対してどのような感情が引き起こされるかを、音声を模倣する確率が異なる3種類のクリーチャを用いて比較評価し、その結果を考察する。

以下、2章では対人的な志向姿勢が誘発される要因について述べ、3章では、反響的に模倣する仮想的なクリーチャとの相互作用に対する心理的な評定実験の結果について述べ、それに対する考察を行う。4章では、本論文のまとめと今後の研究の展開について述べる。

2. 志向的な姿勢の誘発の要因

人間は自分と関わる相手の様々な振舞いからその特性をとらえようとする。Dennettは、人間が関わる対象の性質をとらえる場合の帰属傾向として、i) 物理的な姿勢、ii) 設計的な姿勢、iii) 志向的な姿勢の3つをあげている¹⁾。i) は、自分と関わる対象の振舞いが物理的な法則や物理的な構造に従って表出されたものとしてとらえようとする姿勢である。ii) は、自分と関わる対象の振舞いが設計されたとおりの機能として表出されたものとしてとらえようとする姿勢である。また iii) は、自分と関わる対象の振舞いは背後に存在する何らかの意図や感情に基づいて合理的な判断の下で表出されたものとしてとらえようとする姿勢である。たとえば、人間が遠くにある大きな壁に向かって「ヤッホー」と呼びかけると、しばらくして壁の方から同じように

「ヤッホー」という声が返ってくる。i) では、単純に音が壁に反射して自分の呼びかけがこだましてきたものとしてとらえる。ii) では、壁に何らかの音声の録音/再生装置が組み込まれており、一定時間後に録音された自分の音声再生されたものとしてとらえる。また iii) では、壁が自らの意図に基づいて自律的に呼びかけに応えたものとしてとらえる。

人間は、たとえ関わる対象が機械であっても、その対象の振舞いに対して物理的な姿勢や設計的な姿勢に先立ち、志向的な姿勢でとらえようとする傾向があると考えられている^{5)~7)}。機械との関わりにおいて、ある条件の下では人間の志向的な姿勢が強く引き出される場合がある。たとえば、海外旅行の準備をしている際に必要な情報を表示したコンピュータから、後日ある課題を提示された際に、人間はつい熱心にその課題に取り組んでしまう等、無自覚的に対人的な返報行動を行うことがある^{5),6)}。また、コンピュータ上でインタフェースエージェントと何らかの共同作業を行っている際に、自分の意見に同調したキャラクタが推奨する操作を受容しやすい等、無自覚的に対人的な親和行動を行うこともある⁷⁾。これらの事象より、機械の外観ではなく関わり合う過程で表出される振舞いによっては、人間の機械に対する対人的な反応が無自覚的に引き起こされることが実証されつつある。しかし、人間の志向的な姿勢が引き出されれば、機械との情緒的な関係の形成・維持につながるかどうかは明らかにされていない。本論文では、このような人間の持つ帰属傾向に着目し、志向的な姿勢を引き出す手がかりとなる反響的な模倣音声を表出するシステムが与える心理的な影響を心理評定実験を通して調べていく。

3. 心理的評定実験

本章では、被験者を用いて自分が発した音声をシステムに反響的に模倣された場合に引き起こされる心理的な影響を評価した結果について述べる。

3.1 実験環境

心理的評定実験を行うにあたり、音声を介してやりとりをする対象として、コンピュータ上に仮想的なクリーチャを構築した。このクリーチャは1つの目玉からなるシンプルな外観を持つ。さらに、図1に示すメカニズムに従って、外界から入力された音声を声の大きさ、リズム、イントネーションというプロソディのレベルで反響的に模倣した非分節音を表出する。クリーチャは、人間を含めた既存の動植物の持つ確定的な外観や音声ではなく、多様な解釈が可能な、すなわち不定な外観や音声を備えることで、特別な先入観を与え

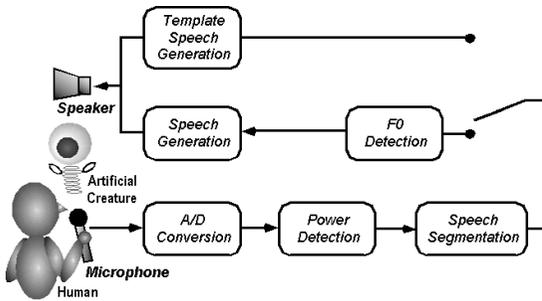


図1 実験システムの構成例

Fig. 1 Example of system implementation for experiment.

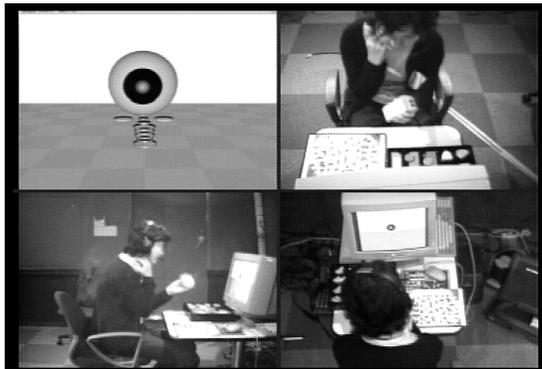


図2 被験者と仮想的なクリーチャとのやりとりの様子

Fig. 2 Example of interactive session between human and artificial creature.

ずに被験者とやりとりをすることが可能となるため、反響的な模倣により引き起こされる心理的な影響を調べるうえで有効なシステムであるといえる(図2)。

システムの処理体系は以下のとおりである。マイクロホンに入力された音声信号を 16 (KHz), 16 (bit) で A/D 変換後 (A/D Conversion), 対数音声パワーを算出し、その時系列パターンを検出する (Power Detection)。その計算結果を基に、閾値を用いて音声区間検出を行う (Speech Segmentation)。音声区間ごとに基本周波数を算出し、その時系列パターンを検出する (F_0 Detection)。本実験では、基本周波数の算出方法として、AMDF (Average Magnitude Differential Function) 法を用いた。音声区間ごとに音声パワーと基本周波数を計算した結果から得られた時系列パターンに基づいて複数のサイン波を組み合わせることにより非分節音に変換する (Speech Generation)。あるいは、あらかじめ用意した音声波形テーブルを参照し、非模倣的な非分節音を生成する (Template Speech Generation)。本実験では、サイン波や三角波等を用いて、音声パワーおよび基本周波数の時系列パ

ターンが一定である 1~4 モーラの非分節音を非模倣的な音声として用意した。クリーチャがこのようなメカニズムを備えることにより、被験者がどのような発話をした場合でも、クリーチャから何らかの応答を得ることが可能となる。

図3(上)に被験者が発した音声(とりさん: torisan)のプロソディをシステムが反響的に模倣した場合、図3(下)に被験者の発した音声(とりさん: torisan)に対してシステムが非模倣的な一定のプロソディの音声を出した場合の音声波形および基本周波数の時系列パターンを各々示す。この2つの場合を比較すると、システムが模倣的な音声を出した場合では、被験者の音声の大きさや高低、リズムの時系列パターンの類似した音声波形と基本周波数の時系列パターンが示されているが、システムが非模倣的な音声を出した場合では、大きさ、高低、リズムのいずれも被験者の発した音声に依らないプロソディー一定の3モーラの音声波形および基本周波数の時系列パターンが示されている。

3.2 仮説

人間は、仮想的なクリーチャが入力された発話を反響的に模倣する機能しか備えていないにもかかわらず、クリーチャの模倣的な振舞いの背後に自分に向かう意図や感情を見いだしてしまう。志向的な姿勢が引き出される結果として、人間はクリーチャから何らかの心理的影響を与えられると考えられる。本実験では、クリーチャの出力する非分節音の内容が模倣的な音声である場合と非模倣的な音声である場合の割合を変化させることによって、反響的な模倣が被験者に与える影響を調べる。

被験者の発話に対して仮想的なクリーチャがつねに非分節音を用いた模倣的な音声のみ、あるいは非模倣的な音声のみを返す場合には、被験者はクリーチャがだれに対しても同様に機械的な応答をしていると解釈する。自分の発話に何らかの応答が返ってくれば、被験者の志向的な姿勢を引き出すことは可能である。しかし、仮想的なクリーチャの出力する応答が模倣的な音声である場合と非模倣的な音声である場合が混在すれば、被験者のクリーチャに対するさらなる志向的な姿勢を引き出す余地が生じる。すなわち、模倣的な音声と非模倣的な音声が入混じる場合、被験者はクリーチャが発する非模倣的な音声を機械的な応答としてではなく、クリーチャの第三者の模倣をしている音声であると解釈することが可能となる。また、同様の場合にクリーチャが反響的に模倣した音声を出力する際には、クリーチャが被験者を個人的に認知した結果として自分の発話を模倣していると見なす。

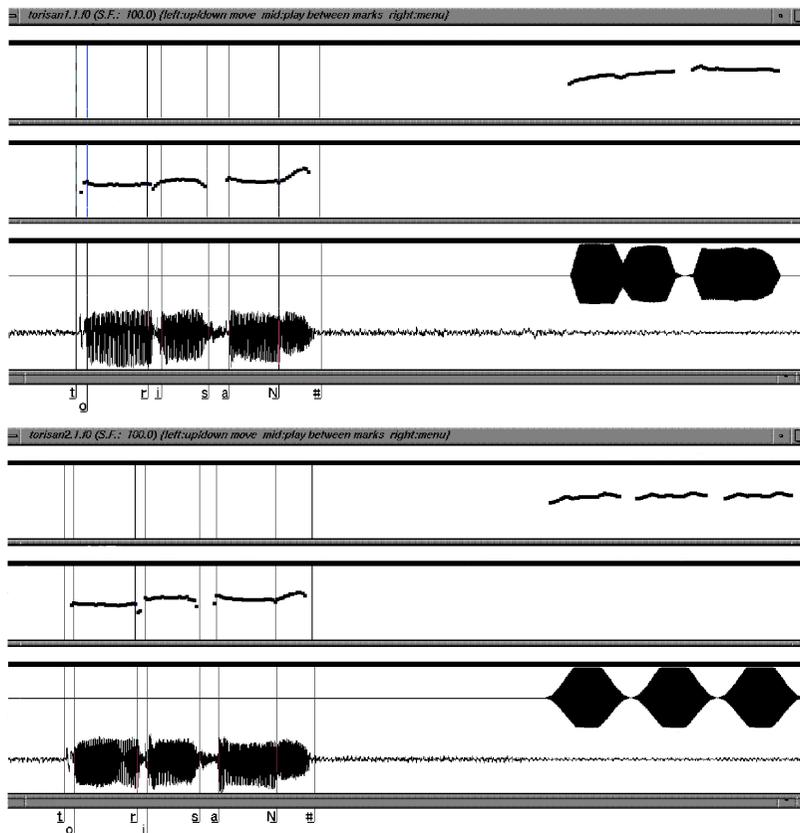


図3 被験者とシステムのやりとりにおける音声波形と基本周波数の時系列パターン：(a) 被験者の発話に対してシステムが反響的に模倣した音声を出力した場合(上)(b) 被験者の発話に対してシステムが非模倣的な音声を出力した場合(下)

Fig. 3 Example of time sequence pattern of sound waves and f_0 on interaction between subject and system: (a) system expresses mimicking voice against subject's utterance (top), (b) system expresses template voice subject's utterance (bottom).

これより本実験では、非分節音を用いて反響的に模倣的な音声を出力する確率が非模倣的な音声を出力する確率よりも高いクリーチャほど、被験者自身を認知したやりとりを行っているという解釈が可能となるため、被験者にポジティブな印象を与えるという仮説の下で心理的評価を行う。

3.3 実験条件

前述の仮説に基づいて、被験者が自分の発した音声を反響的に模倣されることにより引き起こされる心理的影響を調べるために、以下の3種類の音声条件を設定した。

- (a) 仮想的なクリーチャの応答として、直前に入力された発話のプロソディを反響的に模倣した音声である場合が80%、非模倣的な音声である場合が20%の確率で出力される。
- (b) 仮想的なクリーチャの応答として、直前に入力

された発話のプロソディを反響的に模倣した音声である場合が50%、非模倣的な音声である場合が50%の確率で出力される。

- (c) 仮想的なクリーチャの応答として、直前に入力された発話のプロソディを反響的に模倣した音声である場合が20%、非模倣的な音声である場合が80%の確率で出力される。

被験者は、この3種類すべての音声条件の下で仮想的なクリーチャと音声を介して関わり合う。

ただし、仮想的なクリーチャの視線と動作は、音声の条件にかかわらず、次に述べるようなつねに一定の動きを表出するように設定された。視線はつねに被験者のいる方向である正面を向くように設定された。動作は被験者の音声が入力されていない場合には横に揺れているように設定された。また、被験者の音声を知覚している間は動きを停止し、自身が音声を発する場

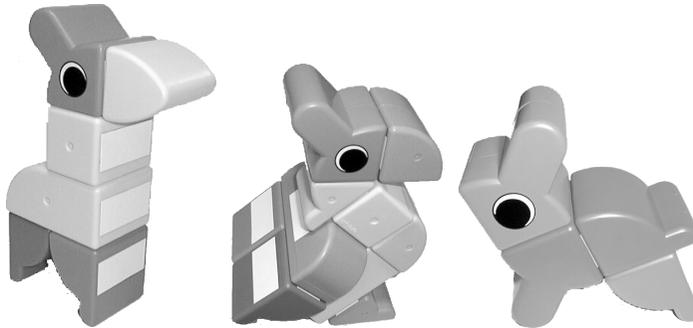


図4 被験者が実験に使用したブロックの組み立て例(きりん(左), あうむ(中), うさぎ(右))
Fig. 4 Example of building blocks by using subject (giraffe (left), parrot (middle), rabbit (right)).

合には通常よりも一回り大きくなるように設定された。

3.4 実験の概要

被験者： 男女の大学生 24 名

手続き： 各被験者に対し、以下の手続きで実験を行った。

- (1) 被験者には、1 分程度の仮想的なクリーチャとのやりとりの例をビデオにて鑑賞してもらう。その間に、実験者からこの実験に関する教示を受ける。
- (2) 教示終了後、実験者は室外に退出し、それと同時に被験者は仮想的なクリーチャとのやりとりを開始する。
- (3) 課題開始後、約 4 分間が経過したところで、実験者は被験者にやりとりの終了を知らせる。
- (4) 課題終了後、被験者は回答用紙において各評価項目に関する質問の回答を行う。
- (5) 他の 2 条件においても同様に、2 ~ 4 までの手続きを繰り返す。ただし、被験者がやりとりするクリーチャの音声条件に関する呈示順序は、順序効果を考慮してカウンターバランスに基づいて決定された。

3.5 教示および評価項目

本実験において、被験者は実験者によって次のような教示を受けた。仮想的なクリーチャは、人間にたとえるならば 1~2 才程度の知能レベルであり、ものの名前を覚えたがる年頃である。被験者は、目前においてあるブロック(図4 参照)を用いて動物や乗り物等の物体を組み立て、その名称を仮想的なクリーチャに向かって繰り返し言うことによって学習させる。

このような課題を与えられることにより、年齢や性別によらず、被験者は仮想的なクリーチャと容易に音声を紹介したやりとりを行うことが可能となる。課題終

了後の心理評定では、人間とロボットあるいは擬人化エージェントとのやりとりにおける親和的印象に関する心理的評定を行った先行研究^{(8),(9)}に基づいて作成した以下の評価項目についての回答を求めた。

(1) 仮想的なクリーチャとのやりとりに対する印象【被験者自身に対する評価】

- 適応度 (adaptability): どの程度仮想的なクリーチャに適応的に接することができたと感じたか。
- 疲労度 (stress-free): どの程度仮想的なクリーチャとのやりとりの過程自体にストレスを感じたか。
- 達成度 (achievement): どの程度課題に対する達成感を感じたか。

【仮想的なクリーチャに対する評価】

- 習得度 (memory retention): どの程度仮想的なクリーチャが被験者の指示した物体の名称を習得できたと思えたか。
- 協調度 (cooperation): どの程度仮想的なクリーチャが協力的な態度であったと思えたか。
- 友好度 (friendliness): どの程度仮想的なクリーチャが友好的な態度であったと思えたか。

(2) 仮想的なクリーチャに対するモダリティごとの印象

- 生物的様相度 (life-likeness): { 視線/動作/音声 } に対してそれぞれ生き物らしさを感じたか。
- 共感度 (sympathy): { 視線/動作/音声 } に対してそれぞれ被験者の気持ちが分かってくれたように感じたか。
- 言語把握度 (verbal understanding): { 視

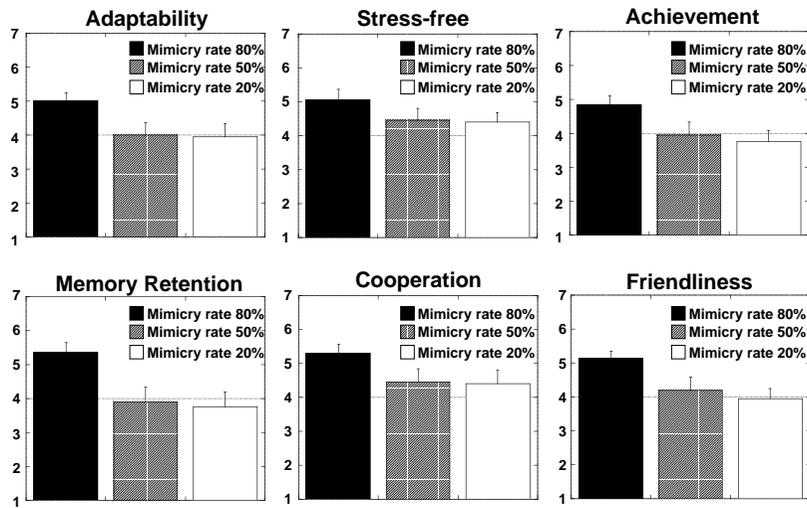


図5 仮想的なクリーチャとのやりとりに関する印象:(a)被験者自身に対する評価(上),
(b)仮想的なクリーチャに対する評価(下)

Fig. 5 Results of evaluation (I): impression of interaction with artificial creature
(a) subjects' attitudes to creature (top), (b) creature's attitudes according to subjects (bottom).

線/動作/音声}に対してそれぞれ被験者の言葉を理解しているように感じたか。

すべての評価項目は、7点尺度で求められた。評価値が大きいほど評価項目の内容を満たしている。

3.6 結果

図5に仮想的なクリーチャとのやりとりに関する印象について、3種類の音声条件における被験者の評価値の平均値を各々示す。また、図6に仮想的なクリーチャの視線/動作/音声の各モダリティに対する印象について、同様に各音声条件における被験者の評価値の平均値を示す。双方の図とも、縦軸に評価値の平均値、横軸に実験条件である仮想的なクリーチャの発した音声のうち、模倣的音声が出力される確率を示す。これらの図は、評価値が4未満の場合は負の方向の評価であり、4以上の場合は正の方向の評価を示す。

ただし、24名中4名の評価値に関しては、被験者の声質によりシステムが模倣的な音声を出力することができなかったため、集計結果に加えなかった。したがって、実験段階ではカウンタバランスに基づいて被験者に各音声条件のクリーチャを提示したが、集計段階では順序効果による影響が完全には相殺されていない値となっている。しかし、図5および図6に示した20名分の平均評価値は、順序効果による影響が完全に相殺されている18名分の値とほぼ同じ値を示し、評価値全体の傾向に対してもまったく変化がない。よって以下では検定力を考慮し、20名分の平均評価値を

用いる。

3.7 考察

各評価項目に対する評価結果に対する考察を以下に述べる。

(1) 仮想的なクリーチャとのやりとりに対する印象

【被験者自身に対する評価】:表1(上)に各項目に対する分散分析の結果を示す。多重比較の結果、適応度に関しては、仮想的なクリーチャが模倣的な音声を発する確率が80%の場合と50%以下の場合の間で統計的に有意な差が見られた。仮想的なクリーチャが被験者の発話を反響的に模倣する確率が50%以下の場合よりも80%の場合の方が被験者に教示で求められた課題を遂行できたという印象をより強く与えたといえる。一方、疲労度に関しては、評価値を見る限り、仮想的なクリーチャが模倣的な音声を発する確率が80%の場合が最も高い値であるが、3種類の音声条件間で統計的に有意な差は見られなかった。仮想生物が反響的に被験者が発した直前の発話の模倣を行う確率の高低は、被験者に与えるストレスに関する印象に大きな影響を与えたわけではないといえる。また達成度に関しては、適応度の場合と同様に仮想的なクリーチャが模倣的な音声を発する確率が80%の場合と50%以下の場合の間で統計的に有意な差が見られた。仮想的なクリーチャが被験者の発話を反響的に模倣する確率が50%以下の場合よりも80%の場合の方が被験者に教示で求められた課題を遂行できたという印象をより強く

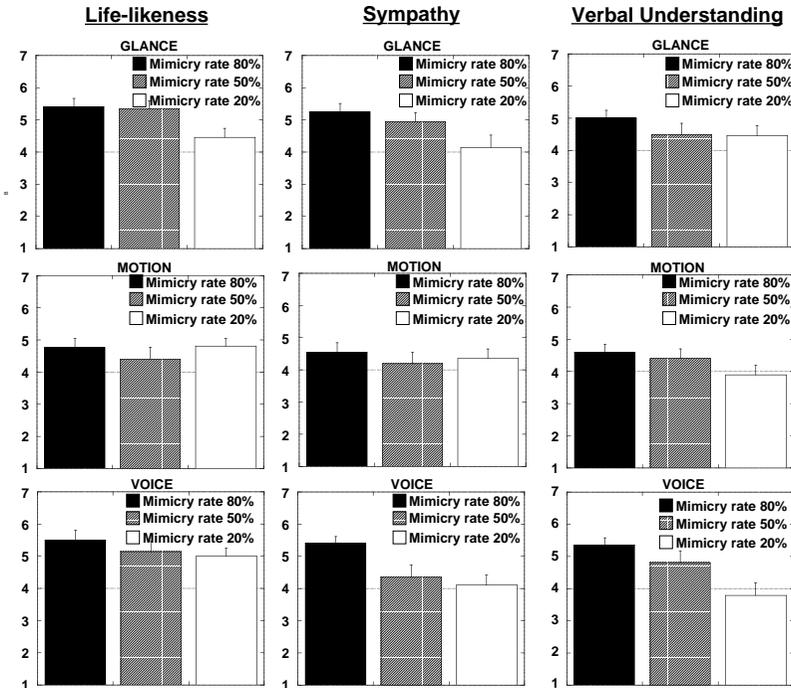


図6 仮想的なクリーチャのモダリティごとの印象:(a)視線(上)(b)動作(中)(c)音声(下)
 Fig.6 Results of evaluation (II): impression of each modality of artificial creature: (a) glance (top), (b) motion (middle), (c) voice (bottom).

表1 仮想的なクリーチャとのやりとりに対する印象における分散分析の結果

Table 1 Results of ANOVA on impression about interaction with artificial creature.

仮想的なクリーチャとのやりとりに対する印象		
被験者自身に対する評価		
適応度	疲労度	達成度
$F_{(2,19)} = 4.496$ $p < .05$	$F_{(2,19)} = 1.519$ $p = .244$	$F_{(2,19)} = 4.693$ $p < .05$
仮想的なクリーチャに対する評価		
習得度	協調度	友好度
$F_{(2,19)} = 7.113$ $p < .05$	$F_{(2,19)} = 2.479$ $p = .111$	$F_{(2,19)} = 5.862$ $p < .05$

与えたといえる。

【仮想的なクリーチャに対する評価】: 表1(下)に各項目に対する分散分析の結果を示す。多重比較の結果、習得度に関しては、仮想的なクリーチャが模倣的な音声を発する確率が80%の場合と50%以下の場合の間で統計的に有意な差が見られた。実際には仮想的なクリーチャが物体の名称を習得していないにもかかわらず、仮想的なクリーチャが被験者の発話を反響的に模倣する確率が50%以下の場合よりも80%の場合の方が物体の名称を覚えたという印象をより強く与えたといえる。一方協調度に関しては、評定値を見る限

り仮想的なクリーチャが発する模倣的な音声の確率が80%の場合が最も高い値であるが、3種類の音声条件の評定値の間で統計的に有意な差は見られなかった。仮想的なクリーチャが反響的に被験者の発した音声の模倣を行う確率の高低は、被験者に与える仮想的なクリーチャの協調的な態度の印象に多大な影響を与えたわけではないといえる。また友好度に関しては、仮想的なクリーチャが模倣的な音声を発する確率が80%の場合と50%以下の場合の間で統計的に有意な差が見られた。仮想的なクリーチャが被験者の発話を反響的に模倣する確率が50%以下の場合よりも80%の場合の方が友好的な態度であったという印象をより強く与えたといえる。

(2) 仮想的なクリーチャに対するモダリティごとの印象【視線】: 表2(上)に仮想的なクリーチャの視線に関する各項目の分散分析の結果を示す。多重比較の結果、どの音声条件の下であっても、一定の視線の動きを表出するようにクリーチャを設計したにもかかわらず、生物的様相度および共感度では統計的に有意な差が算出された。生物的様相度に関しては、仮想的なクリーチャが模倣的な音声を発する確率が50%以上の場合と20%の場合の間で統計的に有意な差が見られた。さらに共感度に関しても、仮想的なクリーチャが模倣的な

表 2 仮想的なクリーチャに対するモダリティごとの印象における分散分析の結果

Table 2 Results of ANOVA on impression about each modality of artificial creature.

仮想的なクリーチャに対するモダリティごとの印象			
	生物的様相度	共感度	言語把握度
視線	$F_{(2,19)} = 6.284$ $p < .05$	$F_{(2,19)} = 4.306$ $p < .05$	$F_{(2,19)} = 1.220$ $p = .317$
動作	$F_{(2,19)} = 2.183$ $p = .521$	$F_{(2,19)} = 3.343$ $p = .557$	$F_{(2,19)} = 1.111$ $p = .206$
音声	$F_{(2,19)} = 1.271$ $p = .303$	$F_{(2,19)} = 5.566$ $p < .05$	$F_{(2,19)} = 5.459$ $p < .05$

音声を発する確率が 80%の場合と 20%の場合の間で同様に統計的に有意な差が見られた。一方言語把握度に関しては、統計的に有意な差は見られなかった。この実験環境における被験者の仮想的なクリーチャの視線に対する印象は、視線の動きのみで与えられるのではなく、音声に対する印象の影響を大きく受けたと考えられる。

【動作】: 表 2 (中) に仮想的なクリーチャの動作に関する各項目の分散分析の結果を示す。多重比較の結果、設定のとおり、動作に関してはいずれの評価項目に関しても、3種類の音声条件の間で統計的に有意な差は見られなかった。このことから、この実験環境においては被験者の仮想的なクリーチャの動作から受ける印象は、音声に対する印象が大きく影響したわけではないといえる。

【音声】: 表 2 (下) に仮想的なクリーチャの音声に関する各項目の分散分析の結果を示す。多重比較の結果、生物的様相度に関しては、どの音声条件においても高い評価値であり (5.0 以上)、条件間に統計的に有意な差は見られなかった。仮想的なクリーチャの反響的な音声の模倣の確率の高低は、仮想的なクリーチャの生物的様相度に対して大きな影響を与えたわけではないといえる。一方、共感度に関しては仮想的なクリーチャが模倣的な音声を発する確率が 80%の場合と 50%以下の場合の間で統計的に有意な差が見られた。仮想的なクリーチャが被験者の発した音声を反響的に模倣する確率が 50%以下の場合よりも 80%の場合の方が被験者に自分の気持ちを理解してもらえたという印象をより強く与えたといえる。さらに言語把握度に関しても、仮想的なクリーチャが模倣的な音声を発する確率が 80%の場合と 20%の場合の間で同様に統計的に有意な差が見られた。仮想的なクリーチャが被験者の発話を反響的に模倣する確率が 20%の場合よりも 80%の場合の方が言語を理解してもらえたという印象をより強く与えたといえる。

3.8 全体考察

仮想的なクリーチャの模倣的な音声を発する確率が異なる 3種類の条件において、統計的な有意差の有無にかかわらずいずれの評価項目に関しても評価値の傾向はほぼ同じであり、80% > 50% > 20%の順に高く評価された。これは、仮想的なクリーチャの発する音声のうち模倣的な音声の占める割合が大きいほど被験者の仮想的なクリーチャに対する印象の評価値も高くなる、すなわちポジティブに評価されたと考えられる。

仮想的なクリーチャとのやりとりに関する印象としては、統計的に有意差が見られた次の 4項目、被験者自身の評価における適応度と達成度、仮想的なクリーチャの評価における習得度と友好度に関して、仮想的なクリーチャの反響的な音声を発する確率が 80%の場合と 50%以下の場合では、評価値が 1ポイント近く異なる結果となった。これより、クリーチャが被験者の音声を反響的に模倣する確率が 80%の場合、クリーチャは被験者個人を認識して模倣していると解釈した結果としてポジティブに評価されたと考えられる。逆に、模倣的な音声を発する確率が 50%以下の場合、被験者に与える印象はほぼ同様であり、自分の発話に対してつねに何らかの応答が返ってきたことが評価されたと考えられる。

また仮想的なクリーチャのモダリティごとの印象としては、どの音声条件下であっても正面方向を向くように設定されていた視線に対して、生物的様相度と共感度に関して模倣的な音声を発する確率が 20%の場合よりも 80%の場合に被験者に対してよりポジティブな印象を与えていた。これより、仮想的なクリーチャに対する印象はモダリティごとではなく、やりとりの過程全体から与えられたと考えられる。

これらの結果より、仮想的なクリーチャの反響的な音声の模倣が表出される確率が高い場合は低い場合と比べてクリーチャの態度の好感度に関する項目が有意に高く評価された。このことから、3.2節で述べたように、被験者は自分の発話を仮想的なクリーチャに模倣されることによって、クリーチャの振舞いの背後に自分に向かう何らかの意図や感情が存在するととらえている可能性が示唆されたといえる。さらに、非分節音を用いて反響的に模倣するという状況設定の下では、被験者はクリーチャの模倣的振舞いを、人間とのやりとりの過程で産出される動物の鳴き声や乳幼児の喃語と同様の原初的なレベルのものとしてとらえやすいため、ネガティブな印象よりもポジティブな印象を持ったと考えられる。逆に、クリーチャが人間の成人と同等の日本語等の既存の言語体系に含まれる音声言語を

用いて反響的な模倣をするという状況設定の下では、被験者はクリーチャに対して人間の成人と同等な自律的な振舞いを期待することが考えられるため、必ずしもポジティブな印象を持つとは限らないといえる。

4. む す び

本論文では、人間が非生物である機械に自分の振舞いを反響的に模倣されることによって引き起こされる心理的影響を調べた。そのために、直前に人間が発した音声のプロソディを反響的に模倣する模倣的音声を発する確率と、人間の発した音声には関係なくプロソディの一定した非模倣的音声を発する確率の異なる3種類の条件を仮定し、各クリーチャとのやりとりに対する印象の比較評価を行った。その結果、以下のような知見を得た。

- (1) 仮想的なクリーチャには何らかの感情や欲求に基づいて模倣を表出する機能はなく、反響的に模倣する機能のみが備わっている。それにもかかわらず、被験者が発した音声を非分節音を用いて模倣する割合が大きくなるにともない、被験者のクリーチャに対する好感度に関する項目がポジティブに評価された。
- (2) 仮想的なクリーチャの視線は、音声の条件によらずつねに正面方向を向くように設計されていたにもかかわらず、生物的様相度と共感性について反響的な模倣を表出する割合が高い場合は低い場合と比べてポジティブに評価された。

これらのことから、被験者は仮想的なクリーチャの反響的な音声の模倣を何らかの感情や欲求に基づいて表出された振舞いととらえた、すなわち反響的に模倣されることを手がかりとして志向的な姿勢が誘発されたと考えられる。さらにその結果として、クリーチャに対するポジティブな感情が引き起こされたことが示唆された。非生物である機械に何らかの感情や欲求に基づいた振舞いを表出する機能が備わっていても、動物の鳴き声や乳幼児の発する喃語に似た非分節音を用いて反響的に発話を模倣するという機能が備わっていれば、自分が機械に個人的に認知されていると解釈することが可能であるため、人間が志向的な姿勢の下で機械と関わってしまう可能性があるということを実証的に示したといえる。ただし、仮想的なクリーチャが日常的に使われている言語を用いて応答した場合に、被験者の志向的な姿勢が引き出されたとしても、クリーチャに対して積極的な感情を持つとは限らない。このような他の状況下で機械に模倣されることによって引き出される心理的影響について調べることが今後

の研究課題としてあげられる。

将来、機械と日常的に接する機会が増えるに従い、人間に与えられる心理的な負担がますます増大していくことが懸念されている。そのためには、人間の心理的特性を利用し、情緒的な関係を形成・維持する機能を備えた機械の設計が、今後、重要な課題になってくる。本論文で得られた知見に基づく応用例の1つとして、教育支援システムがあげられる。反響的に模倣するという機能を備えることから教師としての役割に基づいて関わるという従来の視点からではなく、ユーザである乳幼児や児童と同じ視点に立ち、ともに遊び、ともに学ぶという対等な立場から関わることの可能なシステムの設計が考えられる。さらに、既存の言語体系に含まれない音声である非分節音を表出することによって人間のコミュニケーション欲求を満たすことから、ペットロボット^{10)~12)}やトインタフェース^{13)~15)}をはじめとする実世界で人間と相互作用を行うインタフェースシステムの設計への応用も可能であろう。

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へ感謝いたします。また、システムの実装に尽力していただいた古屋隆志氏、西垣成生氏の両氏に謝意を表します。図3のラベリングをしていただいた栗原美和子さんに感謝します。

参 考 文 献

- 1) Dennett, D.C.: *Kinds of minds*, HarperCollins Publishers (1996). 土屋 俊(訳): *心はどこにあるのか*, 草想社 (1997).
- 2) 鈴木紀子, 竹内勇剛, 石井和夫, 岡田美智男: 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, 情報処理学会論文誌, Vol.40, No.4, pp.1453-1463 (1999).
- 3) 石井和夫, 鈴木紀子, 岡田美智男: 音声インタラクションにおける関係性の獲得, 第12回人工知能学会全国大会, pp.334-337 (1998).
- 4) 石井和夫, 鈴木紀子, 岡田美智男: ユーザにあわせるインタラクションの学習, 第14回ヒューマンインタフェースシンポジウム, pp.83-88 (1998).
- 5) Nass, C., Steuer, J., Henriksen, L. and Dryer, D.C.: *Machines, social attributions, and ethopoia: Performance assessments of computers subsequent to "self-" or "other-" evaluations*, *Human-Computer Studies*, Vol.40, pp.543-559 (1994).
- 6) 竹内勇剛, 片桐恭弘: 人-コンピュータ間の社会的インタラクションとその文化的依存性—互恵

性に基づく対人的反応, 認知科学, Vol.5, No.1, pp.26-37 (1998).

- 7) 竹内勇剛, 片桐恭弘: 人間の判断に同意傾向を示すエージェントに対する社会心理的反応, 信学技報, pp.57-63 (1999).
- 8) 中田 亨, 佐藤知正, 森 武俊, 溝口 博: ロボットの対人行動による親和感の演出, 日本ロボット学会, Vol.15, No.7, pp.1068-1074 (1997).
- 9) Cassel, J. and Thrisson, K.R.: The power of a nod and a glance: envelope vs. emotional feedback in animated conversational agents, *Applied Artificial Intelligence*, Vol.13, pp.519-538 (1999).
- 10) Fujita, M. and Kageyama, K.: An open architecture for robot entertainment, *The 1st International Conference on Autonomous Agents*, pp.435-442 (1997).
- 11) 田島年浩, 斎藤幸弘, 大角雅治, 工藤敏巳, 柴田崇徳: 感情をもったインタラクティブ・ペットロボット, 第16回日本ロボット学会学術講演会予稿集, Vol.1, pp.11-12 (1998).
- 12) 山本浩司, 水谷研治: 高齢者用コミュニケーション支援器具の開発, 日本ロボット工業学会誌, No.131, pp.27-32 (1999).
- 13) Strommen, E.: When the interface is a talking dinosaur: Learning across media with ActiMates Barney, *CHI98*, pp.288-295 (1998).
- 14) Druin, A., Montemayor, J., Hendler, J., McAlister, B., Boltman, A., Fiterman, E., Plaisant, A., Kruskal, A., Olsen, H., Revett, I., Schewenn, T.P., Sumida, L. and Wagner, R.: Designing PETS: A personal electronic teller of stories, *CHI99*, pp.326-329 (1999).
- 15) Johnson, M.P., Wilson, A., Blumberg, B., Kline, C. and Bobick, A.: Sympathetic interfaces: Using a plush toy to direct synthetic characters, *CHI99*, pp.152-158 (1999).

(平成 11 年 11 月 1 日受付)

(平成 12 年 4 月 6 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年同大学院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。1995年より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に関する研究に従事。2000年4月より名古屋大学大学院人間情報学研究所博士課程(後期)に在学中。電子情報通信学会, 日本音響学会, 人工知能学会, 認知科学会各会員。



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所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「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, ヒューマンインタフェース学会各会員。



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。エンタテインメントロボットの開発に従事。人工知能学会, 日本音響学会各会員。



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」の主体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．
