

# ヒトとロボットとの社会的相互行為の組織化 ——随伴性による発話連鎖

藤井 洋之<sup>†</sup> 岡田 美智男<sup>†</sup>

ヒトとロボットの間により自然なコミュニケーションが成立するためには、ヒトがロボットとの間でどのように相互行為を組織化しているのかを分析し、何が相互行為を維持する手がかりとなりうるのかを明らかにする必要がある。本研究では、非分節音を随伴的に返すロボットとの会話データを収集し、相互行為の組織化に対する微視的な視点から分析した。ロボットが発する非分節音は、言語的な情報を含まないが、会話の進行の中では意味を担う。分析の対象としたのは、ロボットに対して色の名前を教えるという3分間の課題場面(9人分)であった。これらのデータから、ヒト-ロボット-ヒトの発話連鎖を単位として抽出し、会話の進行の円滑さを基準にして分類した。分析により、随伴する非分節音の継続時間長が比較的短いものを優先的な応答、継続時間長が比較的長い場合を非優先的な応答とする優先応答体系の形成が観察された。本研究の結果、随伴的に非分節音を返すロボットとのやりとりを繰り返す中での相互行為の組織化の様相の一部が明らかにされた。

## Organizing Human-Robot Social Interaction ——Utterance Sequence Through Contingency

HIROYUKI FUJII<sup>†</sup> and MICHIO OKADA<sup>†</sup>

For mutual human-robot communication, it is necessary to clarify how humans treat robots as partners in social interaction. In this paper, we investigate how humans organize social interaction with an artificial-creature-type robot that utters synthesized inarticulate sound, which doesn't contain linguistic information itself but can play a meaningful role in the flow of interaction. For analysis, the sessions of nine subjects each of whom talking with the robot for three-minute were recorded. These interaction data were analyzed based on the utterance sequences in the flow of conversation. The analysis unit was a specific sequence of three utterances: 1) initiative utterance of the participant, 2) contingent inarticulate sound of the robot, and 3) follow-up utterance of the participant. These utterance sequences were categorized by two types of the interaction flow according to the smoothness of the sequences. That is, some utterance sequence proceed smoothly while the other type contains a kind of trouble and don't proceed smoothly. The analysis results suggest that the duration of a contingent inarticulate utterance of the robot was organized as a kind of the "preference organization" in the flow of interaction.

### 1. はじめに

ロボットは、コミュニケーションの相手となりうると考えられてきた人工物である。近年、現実の社会にコミュニケーションロボットと呼ばれるロボットが登場し、ヒトとロボットとのコミュニケーションや共生、関係形成などが具体的な問題として議論されるようになってきた<sup>1)</sup>。ロボットという存在とのコミュニケーションについて考えることは、そもそもコミュニケーション可能な他者とは何か、他者とのコミュニケーシ

ョンとは何を実現することなのかという、いまだ明らかになっていないコミュニケーションの基本構成について考えることに通ずる。ヒトとロボットとのコミュニケーションや相互行為の研究は、ヒト社会の中での共生を実現するロボットを構築していくという側面に加え、ヒトとヒトとの相互行為の組織化における様相を逆照射することにもなるのである。ロボットとロボット、あるいはヒトとロボットとの相互行為の組織化を実現する中で、ヒトとヒトとの相互行為では自明であるとされていることや暗黙の前提となっている部分を顕在化させることで、新たな理解を得ることが期待される。

本研究は、ヒトとロボットとの相互行為の組織化に

<sup>†</sup> ATR ネットワーク情報学研究所

ATR Network Informatics Laboratories

おける過程を分析する中で、ヒトの社会的な相互行為の組織化の様相を明らかにする構成的なアプローチとして位置付けられる。これまでヒトとヒトとの社会的相互行為は、エスノメソドロジヤや会話分析と呼ばれる社会学の一領域として研究されてきており、ヒトとロボットとの相互行為の組織化を微視的に分析する試みや構成的な視点から明らかにする試みはきわめて新しいと考えられる。

ロボットとのコミュニケーションにおける基本的な問題は、コミュニケーションの基盤となる社会的なつながり (social bond) が形成されないことである。雑談などの日常的なコミュニケーションの成立基盤を考慮すると、情報伝達的なコミュニケーションは、お互いの社会的なつながりを形成・維持するような基盤の上に成り立っていることが分かる<sup>1)</sup>。この基盤は、乳児と養育者の関係という発達の初期段階からはじまり、「原初的なコミュニケーション」と呼ばれる<sup>2)</sup>。本研究では、このような社会的なつながりの形成という原初的な側面に着目して、ヒトとロボットとの相互行為がどのように組織化されるかという様態を明らかにすることを目標としている。

原初的なレベルのコミュニケーションは、相互行為の中で、主に非言語的な媒介によって達成されると考えられる。本研究で用いたロボットは、非分節音を随伴的に発する。同様のシステムを用いて、スクリーン上のエージェントとの間で、社会的関係が形成されることが確かめられている<sup>9)</sup>。非分節音は、動物の鳴き声や乳幼児の喃語のような非言語音声であり、記号的な意味を持たないが、相互行為の中ではその進行に影響する役割を担う。たとえば、前言語期の乳児が発する喃語は、養育者の随伴的な応答を引き出し、養育者との間に社会的な絆をもたらす原初的な相互行為の鍵となっている。本研究では、個々の発話そのものの意味は特定できなくても、相互行為の文脈の中で意味づけられていく、つまり、ある秩序をなして組織化されていく様を分析するために、非分節音を随伴的に発するロボットとの相互行為を観察する。

本研究では、ヒトとロボットの局所的な発話連鎖を単位とした場合の会話の進行を指標として、ロボットの発話がどのように扱われたかについて分析していく。

一般的に、会話という相互行為は、つねに円滑に進行するのではないが、局所的に発生したトラブルが相互行為的な調整によって「修復」されることによって決定的に断絶することなく進行していくと考えられる<sup>3)</sup>。円滑な進行は無標であるのに対して、トラブルは有標であり、これらの相互行為の特徴は外部から観察可能である。

これらの見地から、本研究では、ヒトからの語りかけに対して非分節音で随伴的に反応するロボットを用いて、ヒトとロボットとの相互行為の組織化の過程を微視的に分析する。具体的な分析として、ヒトとロボットとの間で組織化される発話連鎖において、その会話の進行やトラブル、そして修復を行ううえで利用される手がかりとその様式を明らかにするために、実験から得られた発話連鎖データを「ヒトからの語りかけ」「ロボットからの随伴的な応答」「そのロボットの応答に対するヒトの反応」という三つ組で整理し、会話進行の円滑さという観点から分類する。ここでは、ヒトが先行する自分の発話とそれに随伴するロボットの発話の関係をどのように使って会話の進行に反映させているかに着目した。次に、発話連鎖において先行する発話が後続する発話を拘束することに着目し、ヒトの語りかけが後続発話として要求するタイプに対して、随伴的に返されたロボットの発話の適切性をどのように扱うことによって会話を進行させていたかについて分析する。また、会話の中で複数の発話が重複してしまう状況での、ロボットの発話が適切であったかどうかをどのように扱って会話を進行させていたかを分析する。これらのヒトとロボットのインタラクション場面の分析から、ヒトの発話に随伴するロボットの非分節音が、会話の流れの中で意味づけられ、会話という組織を維持していく際リソースとして利用されていることを示す。

以下、2章では、発話連鎖と会話の進行についての分析に対する視点を整理する。3章では、非分節音を随伴的に返すロボットとの相互行為場面の設定と収録について述べる。4章では、その収録データを2章の観点から分析し、ロボットの発話の時間的特徴が会話の組織化のためのリソースとして用いられていることを示す。5章では、分析結果に基づいて、ロボットとの相互行為の組織化について考察する。

## 2. 発話の連鎖による相互行為の組織化

本章では、まず、相互行為としての会話の組織化についての分析に対する視点を整理する。特に、会話の進行におけるトラブルとそれに対する修復現象につい

ここでいう「非分節音 (inarticulated sound)」は、動物の鳴き声のような「単位に分けられない音のかたまり」を意味している。音声の「非分節的 (Supra-segmental)」特徴という用法とは異なり、言語学における、ヒトの音声言語を他の動物の声と区別する特徴としての、メッセージを小さい単位に分けられるという「分節性 (articulateness)」という用法に基づいている (cf. 文献 10))。

て述べる。また、これらの現象を扱う際の分析単位や会話を連鎖させる際の、発話の意味の不確定さについて議論する。

### 2.1 会話の進行における円滑さ

会話は、参加者同士が話し手、聞き手という役割を交替しながら発話を交換することによって進行する。この「相互行為としての会話」がどのように組織化されるのかについては、社会学におけるエスノメソドロジや会話分析 (Conversation Analysis) の立場から、あるいは言語学的な立場からは談話分析 (discourse analysis) として研究が進められてきた<sup>3)-5),7)</sup>。

会話の進行はつねに円滑であるとは限らないが、大局的には何らかの秩序だった相互行為として観察される。話者の交替や発話の連鎖において、局所的に何らかのトラブルが発生した場合でも、参加者である話し手もしくは聞き手によって「修復」され、破綻せずに会話は進行する。このようなトラブルの発生は、会話の参加者だけでなく、観察者にも観察可能である。たとえば、「言い直す」「聞き返す」「沈黙」などの行為は、何らかのトラブルの所在を表示するものである。逆に、会話の進行が円滑である場合は、特定の行為が会話の参加者や観察者にとっての標識となるわけではない。

このように、会話の進行の円滑さという観点から見ると、会話という相互行為は、局所的には齟齬があっても、そのトラブルは参加者間で調整することによって修復され、大局的には秩序だて組織化されるものであると考えられる。

### 2.2 発話連鎖と会話の進行

本研究では、会話の進行に関する分析単位として、発話対とそれに後続する3つめの発話という3連鎖を用いる。発話連鎖は最小構成では発話対から成る。これは、会話分析においては「隣接ペア」と呼ばれる分析単位である。しかし、会話の進行という観点から考えると、2つの連鎖だけでは、流れが観察できない。談話分析において教室での先生と生徒の会話を分析するときなどに用いられる、「IRF (Initiation-Response-Follow-up)」という3つの発話連鎖から成る発話交換構造のように、第3の発話行為が会話の進行を表示する。

発話対において、先行する発話は、後続する発話を「要求」する。先行する第1の発話は、後続する発話を予定して繰り出されたものであり、その意味で後続する第2の発話がどのようなものかは先行する発話に拘束されている。「質問-応答」というような隣接ペアでは、先行発話である「質問」は後続する発話として

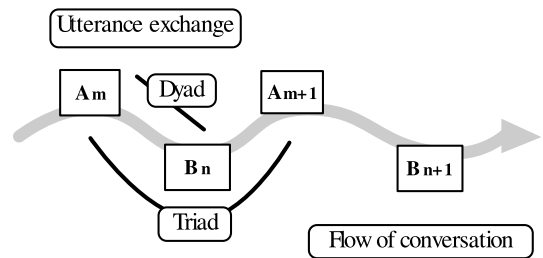


図1 発話交換と会話の進行

Fig. 1 Utterance exchange and the flow of conversation.

「応答」を要求するといえる。つまり、先行する第1の発話には、「後続する発話としてどのような発話を要求するか」が表示されている。

先行発話の要求と後続発話が喰い違った場合は、そこに何らかのトラブルが発生する。しかし、トラブルの発生は、第2の発話までで表示されるわけではなく、発話対に続く第3番目の発話もしくは行為を観察してはじめて特定される<sup>8)</sup>。このように考えると、発話対と、それに後続する発話から成る3つの発話連鎖が、会話の進行に関わる発話連鎖を分析するうえで、最小の単位となる(図1)。

## 3. ロボットとの相互行為データの収集

### 3.1 目的

ヒトとロボットとの相互行為の組織化において、ヒトの発話に随伴して発せられるロボットの非分節音がどのように利用されて発話連鎖が組織化されるかを明らかにするために、防音の施された実験室においてヒトとロボットとの相互行為場面をビデオに録画した。より自然な相互行為を観察するために、具体的な会話方法については例示せず、被験者が自発的に会話を行うようにした。ただし、語りかけるキッカケを与えるために、ロボットに対して「色の名前を教えてあげる」という課題を設定した。

### 3.2 方法

コミュニケーションロボット：本実験に使用したコミュニケーションロボット (Muu) は、音声認識・合成、姿勢や身体配置、視線の移動に基づく社会的表示、ヒトとの間の距離の調整など、ヒトとの社会的相互行為のための機能を備えている<sup>1)</sup>。外見のデザインは、大きな目玉を特徴とした、高さ約40cmの涙滴型である(図2)。表面は発泡ウレタンで作られており、柔らかい。

今回の収録で用いた設定では、Muuの動作は、対面者の音声による語りかけに対して非分節音を返す反響模倣システムを基本としている。それ以外では、下



図2 コミュニケーションロボット Muu  
Fig. 2 Muu: Embodied artificial creature.



図3 実験セッションの例  
Fig. 3 Example of the interaction session.

部外周の距離センサを利用して、わずかに前後に動くことによって対面者との距離を保つようになっている。Muu に搭載された反響模倣システムは、対面者の発話音声を、Muu の眼部の上部に内蔵されたマイクからリアルタイムに取り込み、その入力音声のパワーの時系列パターンを抽出し、複合サイン波と重畳した合成音をスピーカから返すシステムである。このシステムは、対面者の音声を基に一定ではない非分節音が合成され、かつ、対面者の語りかけに対して随伴的に応答することから、原初的な対話システムとして使われてきている<sup>11)</sup>。実際のやりとりでは、約2割の割合で入力音声に基づかない内蔵の非分節音を発する仕様であったこと、そして、マイクとヒトの口元の距離が離れていたために音声区間検出システムが必ずしも働かなかったことから、字義どおりヒトの語りかけを「反響的に模倣する」というよりは、発話開始タイミングや継続時間長などにある程度の幅がある「返事」をするような特徴を持っている。

実験参加者：19歳～23歳の大学生，9名分（男性4名，女性5名）のデータを用いた。参加者は「ロボットとのインタラクション実験」という名目で集められた。事前のアンケートによると、すべての参加者は、今回用いた Muu はいままでまったく見たことも聞いたこともなく、事前の予備知識はまったくなかった。

手続き：参加者個別に、防音スタジオ内で、机の上に乗った Muu の前に座る。その側方から、デジタルビデオカメラによって撮影・録音した。

各参加者は、3分間のインタラクションを1セッションとし、2回の練習と合わせて計4セッションを行った。

まず、練習セッションとして3分間のセッションを2回行った。「Muu は動物の鳴き声もしくは乳児の喃

語のような言葉を話します。3分間、話し相手になって仲良くなってください」という指示を行い、何について話すかは特に指定しなかった。

続く実験セッションでは、被験者に数色の色のついたクッションを渡し、「Muu に色の名前を教えてあげてください」という指示を追加して、同じく3分間のセッションを2回行った（図3）。

#### 4. データ分析

収録した実験セッションでは、すべての参加者が、子どもの相手をするときのような発話や姿勢・態度・指示などで振る舞うことが観察された。

詳細な分析のために、各発話を単位とした局所的な連鎖を切り出して、会話の進行に対する影響を基準として、ヒトの語りかけに対するロボットの発話開始タイミングや発話の重複、ロボットの発話の継続時間長という時間的特徴との関連について考察する。分析では、参加者1名について1セッション（ロボット側の動作不良があった場合を除いて、最初の実験セッション）を対象とした。

##### 4.1 分類I：発話連鎖の抽出

まず、分析の単位となる「発話」を、ヒト・ロボットの発話者ごとに、録音データから抽出した。同一話者の1発話は、400ms以上の無音区間によって区切られたものとした。それぞれ、発話の開始時間、終了時間、継続時間および、間隔を記録した。

次に、これらの発話を単位として、ヒトの先行発話（H1）にロボットが後続し（H1とMが重複している場合を含む）、さらに第3の発話としてヒトの発話が後続するという発話連鎖（H1-M-H2連鎖）を抽出した。

表1は、それぞれのセッションにおける参加者の総発話数とロボットの総発話数、および発話連鎖の数

実際、今回のインタラクションの事後分析では、ヒトの語りかけと反響模倣した合成音との継続時間長の相関係数を計算したところ、全体では  $r=0.19$  ととても低いものであった。

表 1 発話数と発話連鎖数  
Table 1 Number of utterance and Utterance sequence.

Subject	H-ALL	M-ALL	H1-M-H2 sequence	H1-M Overlap
S1	57	25	16	4
S2	67	49	46	8
S3	76	35	35	12
S4	43	21	15	1
S5	30	17	15	5
S6	43	39	35	4
S7	57	40	39	2
S8	64	38	35	9
S9	51	18	17	2
Sum	488	282	253	47
Average	54.2	31.3	28.1	5.2

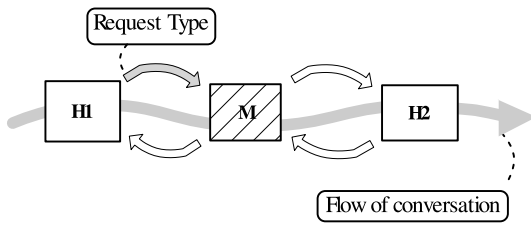


図 4 発話連鎖と会話の進行

Fig. 4 Utterance sequence and the flow of conversation.

である。H-ALL はヒトの発話総数，M-ALL はロボットの発した非分節音の総発話数を表す。また，H1-M Overlap は，ヒトとロボットの発話が重複箇所の数を示している。Muu の発話数と連鎖の数を比較すると，Muu の発話のほとんど（約 9 割）が H1-M-H2 連鎖となっているという参加者間の共通の傾向が見てとれる。よって，以下では，H1-M-H2 連鎖について，参加者全体のものをもとに扱って扱うものとする。

4.2 分類 II：会話の進行に基づく発話連鎖の分類

2 章の会話の進行と発話連鎖のモデルに基づいて，以下の節で示す基準に各々従って，第 1 著者と実験の詳細については知らされていない協力者の 2 名が，H1-M-H2 連鎖の会話の進行タイプと，H1 の要求タイプのタグ付けを行った（図 4）。

4.2.1 H1-M-H2 連鎖の会話の進行タイプ

2 章のモデルに従えば，会話が円滑に進行する場面では，相互行為の当事者も観察者にとっても，会話の進行の円滑さは特に意識されることはない。しかし，会話が円滑に進行していない場合は，何らかのトラブルが観察される。トラブルの所在は，当事者の後続する行為，たとえば，後続する発話における，言い直しや言い淀み，沈黙などとして，観察者にとっても明示される。このような観点に基づいて，前出の 2 名が，会話の進行が円滑かどうかを基準として，H1-M-H2

連鎖を以下の 3 つのカテゴリに分類した。

- Nt (No trouble): H1-M-H2 連鎖において，会話が円滑に進行している [H1「これは赤。分かった？」M「(非分節音)」H2「じゃあ，これは？」]。
- T (Trouble): H1-M-H2 連鎖において，会話が円滑には進行していない(問い直しや沈黙があった場合)[例：「この色は何？」M「...(沈黙 3 秒)(非分節音)」H2「この色は何？」]。
- misc: どちらともいえない。

分類を行った 2 名の一致率 (Cohen’s  $\kappa$ ) は 0.86 であった。これは十分高い一致率であり，分類項目は妥当であったと考えられる。以下の分析では，分類が比較的曖昧である misc を除き，Nt と T の項目のうちで，さらに分類者間で一致していた項目のみを対象とした。

4.2.2 H1 の要求タイプ

次に，ヒトの先行する語りかけ (H1) について，後続する発話としてどのような種類の反応を要求するかという観点から，以下の 4 つの基準で分類した。

- dR (demand Response): 応答要求。Yes/No 質問や呼びかけのように，日本語でいう「はい」や「うん」や「いいえ」などで答えられるような先行発話 [例：「これはあかいですか？」「分かった？」]。
- dA (demand Answer): 回答要求。日本語でいえば，単に「はい」「いいえ」では適合せず，具体的な単語を含んだ発話を要求するような先行発話 [例：「これはなににいますか？」「どうしたの？」]。
- dRA (demand Response/Answer): 後続する発話として，応答・回答のどちらもあてはまる [例：「あかいろ (同意を要求する確認，復唱要求のどちらとしてもありうる)」。]
- misc: どれにもあてはまらない [例：「じゃ (言いかけているが間を空けている)」「そうそう (相手の返事を要求しない)」。]

分類を行った 2 名の一致率 (Cohen’s  $\kappa$ ) は 0.82 であった。これも，十分高い一致率であり，分類項目は妥当であったと考えられる。以下の分析では，分類が曖昧であるものを取り除くため，分類基準が比較的明確ではない dRA と misc は除外した。さらに，dR と dA のうち，分類者間で不一致だった項目を除外し，分類者間で一致していた項目のみを対象とした

一致率は，カテゴリ全体について計算した。Cohen’s  $\kappa$  は次式で得られる。 $\kappa = (Po - Pc) / (1 - Pc)$ 。ただし，Po: 観察された単純一致率。Pc: 偶然一致する確率。

表 2 H1 の要求タイプと会話の進行タイプ

Table 2 The demand type of H1 and the flow type of H1-M-H2 sequence.

Demand type	Flow type	Number	Total
dR	Nt	42	68
	T	19	
	misc	7	
dA	Nt	60	105
	T	32	
	misc	13	
dRA	Nt	25	39
	T	10	
	misc	4	
misc	Nt	30	41
	T	10	
	misc	1	
			253

4.2.3 発話連鎖：H1 の要求タイプと会話の進行タイプ

前 2 項の分類によって、抽出された H1-M-H2 連鎖それぞれについて、H1 の要求タイプおよび、会話の進行タイプのタグが付けられた。以下の分析は、これらのタグを元に行う。H1-M-H2 連鎖における H1 の要求タイプおよび、会話の進行タイプとして分類された事例数の内訳を表 2 に示す。

4.3 分析 I：先行発話の要求タイプと会話の進行ヒトの先行発話に後続するロボットの非分節音による発話行為の随伴性を、後続する非分節音の継続時間長から整理する。

H1-M-H2 連鎖は、要求タイプ (dR/dA) と会話の進行タイプ (Nt/T) で分類すると、以下の 4 つの組合せとなる。

- A1：H1(dR)-M-H2(Nt)
- A2：H1(dR)-M-H2(T)
- B1：H1(dA)-M-H2(Nt)
- B2：H1(dA)-M-H2(T)

この 4 つの組合せにおいて、H1 に随伴する非分節音 M の継続時間長の分布を見る。具体的には、同一要求タイプ内で、Nt になる場合と T になる場合の、M の継続時間長の分布が異なる傾向となるかどうかを検定する (A1 と A2、および、B1 と B2 の比較)。継続時間長の分布には正規性と等分散性を仮定せず、ノンパラメトリックな検定方法である Mann-Whitney の U 検定を用いて、H1-M-H2 連鎖が Nt になる場合と T になる場合で非分節音 M の継続時間長 (M-Duration) の分布に差があるかどうかを計算した。それぞれの頻度分布 (%換算) を図 5 に示す。

U 検定の結果、A1 と A2 を比較すると、A1 における M-Duration の分布は、A2 のそれよりも低い方向

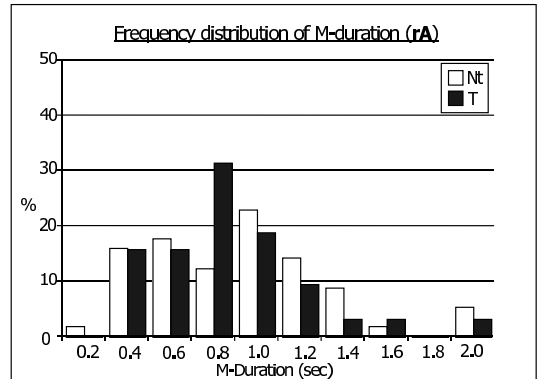
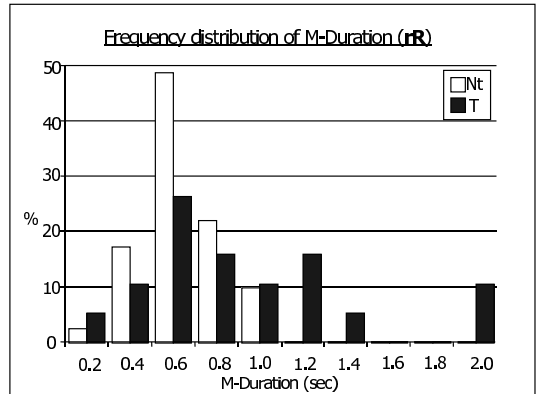


図 5 会話の進行タイプと非分節音の継続時間長の頻度分布  
Fig. 5 Frequency distribution of M-duration and flow type of interaction.

にずれるという有意な傾向が観察された ( $|z| = 2.33 > 1.96, p < .05$ )。この結果は、H1 が応答要求 (dR) の場合、M の継続時間長が比較的短いときはトラブルとならず、円滑に会話が行進する傾向があることを意味する。逆に、非分節音の継続時間長が比較的長いときはトラブルとなる傾向があることを意味する。

この理由として、非分節音の継続時間長が比較的短い場合が「応答」と見なされる場合が多いという傾向は、日本語で応答として用いられる単語が「はい」や「いいえ」などの少ない短い単語に相当するという性質が利用されていることが考えられる。

B1 と B2 の比較では、M-Duration の分布には差が見られなかった (U 検定,  $|z| = 0.93 < 1.96$ )。これは、回答要求では、非分節音の継続時間長の影響が見られなかったことを示している。この理由としては、本実験における回答となりうる非分節音としては、日本語の色の名前に相当するものが多く含まれると考えられるが、日本語の色の名前は「むらさきいろ」「みどりいろ」などの発話するのに約 1 秒かかるような比較的長い単語から「あか」や「あお」「きいろ」などの

約 200 ms 程度で発話できる比較的短い単語までが存在するため、非分節音の継続時間長の長短が回答かどうかを分ける情報にならなかったことが考えられる。また、そもそも一般的な日本語による会話を考えた場合、回答に相当する言葉は、応答と比較すると多様であり、その継続時間長も多様であることも影響している可能性がある。

#### 4.4 分析 II：重複発話事態における会話の進行

本節では、ヒトとロボットの発話が時間的に重複した事態において非分節音の継続時間長が会話の進行にどのように影響を及ぼしたかを分析する。ヒト同士の発話対では、後続する発話タイミングは、会話の進行の円滑さに大きく影響する。たとえば、後続する発話が早すぎる場合は、二者の発話が重複する事態となり、発話対が連鎖せず、トラブルが発生する可能性が高い。今回の分析データの T の 71 事例（表 2 の 4 カテゴリの T の総計）のうち約半数の 34 事例では、先行発話 H1 の終了時点から後続するロボットの発話開始時点までの「発話開始タイミング」が 0 以下であり、重複発話事態であった。

しかし、重複発話事態であれば必ずトラブルとなるとは限らない。ヒト同士の会話では、基本的にはある時点での話者は 1 人であるという話者交替のルールは守られるが、同時に発話が起きることも少なくない<sup>12)</sup>。双方が同時に発話してしまった場合は、互いに相手の発話を妨害したとして、そこにトラブルがあったと見なされる場合がある。一方、発話が重複しても、片方があいづちと見なされる発話である場合などにおいては、会話は円滑に進行すると考えられる。

本節では、ヒトとロボットの発話が時間的に重複した事態において非分節音の継続時間長が会話の進行にどのように影響を及ぼしたかを分析するために、発話連鎖が時間的に重複した事態について、会話の進行のタイプ別に非分節音の継続時間長を比較する。

H1 に対して M の開始タイミングが 0.1 秒以下の場合を「重複発話事態」とした（表 3）。重複発話事態のうち、会話がそのまま円滑に進行する場合（H1-M-H2(Nt)）と、会話の進行に何らかのトラブルがあったと見なされた場合（H1-M-H2(T)）の各々における、非分節音 M の継続時間長（M-Duration）の分布を比較した（図 6）。

- C1：H1-M-H2(Nt)
- C2：H1-M-H2(T)

U 検定の結果、C1 < C2 という M-Duration の分布の差は有意となった（ $|z| = 1.72 > 1.68$ ,  $p < .05$ ）。これは、重複発話事態において会話の進行にトラブル

表 3 重複発話事態と会話の進行タイプ  
Table 3 Overlap and the flow type of interaction.

Flow type of H1-M-H2	Number
Nt	23
T	36
misc	5
sum	64

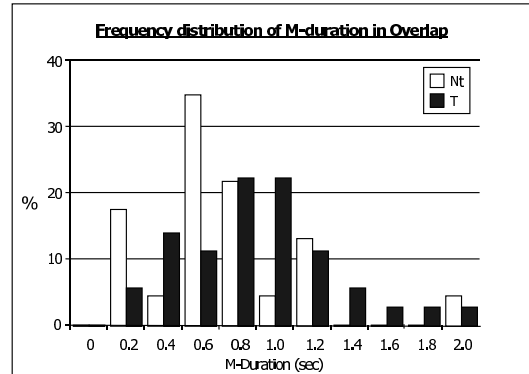


図 6 重複発話事態における非分節音の継続時間長の頻度分布  
Fig. 6 Frequency distribution of M-duration in the overlap condition.

があった場合となかった場合では、随伴する非分節音の継続時間長の分布が異なっていたことを意味している。C1 における M-Duration が C2 のそれよりも有意に短かったということは、非分節音 M の継続時間長が比較的短い場合はトラブルとならず会話がそのまま円滑に進行する傾向があり、逆に、M の継続時間長が比較的長い場合はトラブルとなる傾向があったことを示す。この結果は、重複発話事態において、短い非分節音は「あいづち」のような役割として扱われ、逆に、長い非分節音は会話の進行を妨害する発話であるとして扱われたことを示唆する。

## 5. 総合考察

本研究では、参加者が非分節音を随伴的に発するロボット Muu と実験室内で短時間の会話をする相互行為場面を対象に分析を行った。総じて参加者は、子どもに相対するときのような態度で、ロボットとある程度自然な会話を行っていた。参加者の発話と随伴するロボットの発話の発話連鎖を、会話の進行の円滑さという観点から分析した。その中で、随伴するロボットの発話の継続時間長が会話の進行に影響を与えていたことが明らかとなった。

まず、先行発話の要求タイプと会話の進行における非分節音の継続時間長の分布の分析（分析 I）では、ヒトの先行発話のタイプが応答要求の場合、随伴す

るロボットの非分節音の継続時間長が短いと会話は円滑に進行し、長いと会話の進行に乱れが生じるというパターンが統計的傾向として観察された。このようなパターンは、「優先応答体系（選好的応答による組織化）<sup>13),14)</sup>」に相当する組織化を示していると考えられる。たとえば、我々の会話では、先行発話の問いに対して肯定の場合はすぐに答える（優先的応答）のに対して、否定の場合は答えを発するまでの間が長い（非優先的応答）というような優先応答体系が観察される<sup>14)</sup>。この分析結果は、随伴的に非分節音を返すロボットとのやりとりにおいても、非分節音の継続時間長の長短を組み込んだ優先応答体系という、日常的なヒト同士の会話と共通するパターンの組織化が観察されたことを意味している。

分析 II では、重複発話事態での会話の進行における非分節音の継続時間長の分布を比較し、あいつち等の会話の流れを妨害しない発話となるか、会話の流れを妨害する発話となるかは、随伴する非分節音の継続時間長に影響されていることが示された。この結果もまた、ヒトの先行発話に重複する非分節音の継続時間長が比較的短いものが優先的応答、継続時間長が比較的長いものが非優先的応答、という優先応答体系を形成するというかたちで会話が組織化されていたことを示している。

従来の会話分析の知見では、相互行為の組織化の様相は、相互行為の当事者の社会文化的背景に依存するため、ヒトとロボットという文化が共通するとは考えられない者の間の相互行為で同様の方法論で分析が可能か、組織化の様相を見出すことができるかどうかは明らかではなかった。本研究の分析の結果、ヒトとロボットの相互行為においても、表面的な様相は異なっているが、相互行為の組織化の構造については分析が可能であることが示された。

また、応答に非分節音以外の言語音を用いた場合には応答を継続時間長だけで比較することはできないため、従来の研究では、応答の継続時間長が会話を円滑に進行させるかどうかに関わるという事実は明らかにされてこなかった。この結果は、音声対話によるシステムの操作などに応用可能であると考えられる。

今回の分析では、データをなるべく定量的に分析するために、比較的似かよった層（20歳前後の大学生）とロボットとの相互行為を分析の対象とした。今後、性差・年齢差・文化差を含む個人性に基づく質的分析を適用することによって、相互行為の組織化に関わる新たな知見が得られると考えられる。

また、今回の分析では、非分節音の特徴として継続

時間長のみを対象としたが、ほかに、非分節音の韻律的特徴や、発話タイミングの時間的調整などの要因も相互行為の組織化における手がかりとして利用されていると考えられる。これらの要因に関しても、会話の進行の円滑さに与える影響という観点から、同様の分析が可能であろう。また、音声だけではなく、視線方向や身体配置なども、相互行為の組織化に利用されていると考えられる。これらの時間的・空間的な微視的な調整過程について、相互行為の組織化という観点から分析することが、ヒトとロボットとの相互行為を明らかにするうえで必要である。

これらのロボットとの相互行為においてヒトが利用している手がかりを、ロボットもまた利用可能になることが、ヒトとロボットの相互行為が対等なたちで成立するための条件となる。ロボットが、会話の進行の円滑さやトラブルを感知し、随伴する行為を自ら調整するシステムを実装することによって、ロボットが相互行為の組織化に寄与することができるようになっていくと考えられる。発話を上手に連鎖させていくというよりはむしろ、会話の進行におけるトラブルを発見し、修復することを備える必要がある。このようなヒト同士の相互行為では基盤となっていて特別視されない相互行為のプロセスは、ヒトとロボットとの間の相互行為を分析していくことで明らかになっていく側面があると考えられる。

謝辞 本研究に関して有益な議論をしていただいた Muu Project のメンバに感謝いたします。

本研究は独立行政法人情報通信研究機構（NiCT）の研究委託「人間情報コミュニケーションの研究開発」により実施したものである。

## 参 考 文 献

- 1) 岡田美智男：社会的な相互行為とそのリアリティを支えるもの、身体性とコンピュータ、岡田美智男、三嶋博之、佐々木正人（編）、pp.220-232、共立出版（2000）。
- 2) 鯨岡 峻：原初的コミュニケーションの諸相、ミネルヴァ書房（1997）。
- 3) 高梨克也：社会的相互行為を「見る」方法、人工知能学会誌、No.16, Vol.6, pp.799-805（2001）。
- 4) Sacks, H., Schegloff, E.A. and Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation, *Language*, Vol.50, No.4, pp.696-735（1974）。
- 5) Goodwin, C.: *Conversational organization: Interaction between speakers and hearers*, Academic Press, New York（1981）。
- 6) Coulthard, M. and Brazil, D.: Exchange struc-



ture, *Advances in Spoken Discourse Analysis*, Coulthard, M. (Ed.), pp.50-78, Routledge (1992).

- 7) 石崎雅人, 伝 康晴: 対話の構造, 談話と対話, 石崎雅人・伝 康晴 (編), pp.135-175, 東京大学出版会 (2001).
- 8) Schegloff, E.A.: Third Turn Repair, Towards a Social Science of Language: Papers in honor of William Labov. Volume 2: Social Interaction and Discourse Structures, Guy, G.R., Feagin, C., Schiffman, D., and Baugh, J. (Eds.), pp.31-40, John Benjamins, Amsterdam: (1997).
- 9) 鈴木紀子, 竹内勇剛, 石井和夫, 岡田美智男: 非分節音による反響的な模倣とその心理的な影響, 情報処理学会論文誌, Vol.41, No.5, pp.1328-1337 (2000).
- 10) 辻 幸夫 (編): ことばの認知科学事典, p.87, 大修館書店 (2001).
- 11) 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・インターフェース・シンポジウム, pp.83-88 (1998).
- 12) Schegloff, E.A.: Overlapping talk and the organization of turn-taking for conversation, *Language in Society*, 29, pp.1-63 (2000).
- 13) 高梨克也: 社会的相互行為を見る方法: エソノメソドロロジーと会話分析の基礎, 人工知能学会研究会資料 SIG-SLUD-A002-6, pp.29-36 (2000).
- 14) Pomerantz, A.: Agreeing and disagreeing with assessment: some features of preferred/dispreferred turn shapes, *Structure of Social Action: Studies in Conversation Analysis*, Atkinson, J.M. and Heritage, J. (Eds.), pp.57-101, Cambridge University Press (1984).

(平成 16 年 5 月 18 日受付)

(平成 17 年 1 月 7 日採録)



藤井 洋之

1994 年京都大学文学部 (心理学教室) 卒業. 1996 年名古屋大学大学院人間情報学研究所修士課程修了, 1999 年同研究科博士後期課程満了. 同年より名古屋大 COE 統合音響情報研究拠点特別研究員, 2002 年より ATR 知能ロボティクス研究所等を経て, 2004 年より ATR ネットワーク情報科学研究所研究員. 日常的な状況におけるコミュニケーションの成立基盤や, 音声や音の知覚システムに興味を持つ. 日本認知科学会, 日本音響学会各会員.



岡田美智男 (正会員)

1960 年生. 1981 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. 工学博士. 同年 4 月 NTT 基礎研究所勤務, 1995 年 3 月 ATR 知能映像通信研究所, 2001 年 10 月 ATR メディア情報科学研究所等を経て, 現在, ATR ネットワーク情報科学研究所生態学的コミュニケーション研究室室長. 1998 年より京都大学大学院情報学研究科 (連携講座) 客員助教授を兼務. 社会的相互行為, 社会的ロボティクス, 原初的コミュニケーション等の研究に従事. 日本生態心理学会, 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.