

説得性の社会的構成に着目した  
メディア技術とその応用に関する研究

2014年3月

博士（工学）

大島直樹

豊橋技術科学大学



---

## 説得性の社会的構成に着目したメディア技術とその応用に関する研究

### 論文要旨

近年、人との親和的な関わりを可能とする情報処理システムの実現に向けて、ソーシャルなロボットやエージェントへの関心が高まりつつある。特に、ユーザーの行動を思わず揺り動かすことの可能な説得性を有するロボットとのコミュニケーション技術に関する研究は、パースエイブ・ロボティクスと呼ばれ、その研究開発が精力的に行われている。本論文では、説得性の社会的構成という新たな観点を導入し、人とメディアやロボットとの説得性を伴うコミュニケーション技術の提案と応用分野の開拓を目的としている。

本論文は、7章から構成されている。第1章は、本論文の諸言であり、研究背景と目的を述べている。第2章は、生態心理学におけるオプティカルフローフィールドを社会的インタラクションの場に拡張することで、会話参加者の行動を揺り動かすことの可能なメディア技術を提案している。第3章では、ネットワークを介した複数のコミュニティ間に、第2章で提案した手法を展開し、その有効性を検証している。第4章では、エスノメソドロジーの手法によりインタラクションの詳細を分析し、会話参加者の行動を揺り動かす説得性は社会的に構成されることを示している。

以上の第2章、第3章、第4章における実システムを用いた実験結果に基づく分析により、人とメディアやロボットとの説得性を伴うコミュニケーション技術の提案を行った。本研究では、人とロボット/ソフトウェアエージェントとのインタラクションにおける応用例を示すことも、本研究の目的を実現するうえで重要な役割を果たしている。そこで、第5章では、以上の知見を応用し、Sociable Spotlightと呼ぶソフトウェア・エージェントを設計し、その有効性をフィールド実験により検証している。第6章では、聞き手とのカップリングに基づく発話生成システムを実現し、説得性を伴うロボットのコミュニケーション技術としての有効性を検証している。第7章は、結言である。

## **The Impact of Interactive Artifacts to Ground the Social Interactions and its Applications**

### Abstract

The contents of this thesis is containing three interactive media (Table Talk Plus, Nexus+ and Sociable Spotlight) and two social agents (Sociable Spotlight designed with CG animation and Talking-Ally) for exploring how participants utilize those interactive artifacts and are persuaded by them in human-human/human-robot social interaction.

Table Talk Plus (Chapter 2) is an artifact for promoting mutuality and social bonding among dialogue participants. As it is widely known, spoken utterances can be regarded as an ‘artifact’ that constrains and drives our utterance and thought. In our everyday conversation, spoken utterances and social displays of dialogue participants become important ‘artifact’ as a resource to organize conversational sequences in talk-in-interaction. In this chapter, we are focusing on how to visualize “conversational field” as an artifact in organizing multi-party conversation. So our research question here is how to model and visualize the conversational field that is spontaneously emerged from interactions between dialogue participants. And we propose Table Talk Plus system, which can displays conversational dynamics in a participation frame using simple CG animations. And some research findings from dialogue experiments with the Table Talk Plus are discussed.

Nexus+ (Chapter 3) is an interface for creating field of promoted social interaction on the network. In general, we have three kind of interaction which motivates to swing their body: “physical interaction” which moves the body by applying physical power from someone, “informational interaction” elicits to do their actions from useful information, including language etc., and there is an interaction through a ‘field’ which is people’s action is urged involuntarily according to context. Reed is exploring “field of promoted action” according to the Ecological Psychology, but still difficult to find out how it’s evoke or born in social interac-

tions among the people. We have motivated to develop artificial field to empathize the “field of promoted action” in social interaction on the network among the users and also investigate the technology and how to utilize (handle and promote) it in “field of promoted action”.

Sociable Spotlight (Chapter 4 and Chapter 5) is cognitive artifact to enhance engagement in conversation. The idea is here that understanding why people treat simple geometric animations like real agent which has intention to interact with people even if its geometry is artificial thing will aid the ‘agency’ problems of human-agent interaction. In this chapter, we explore effects of treating simple geometric animations as a real participant to facilitate multi-party conversation in social interaction. Observational study was conducted with groups of two or three persons using simple circle (sociable spotlight) which moves based on dynamic information in the current multi-party conversation, with the goal of discovering how participants are utilizing the behaviors of sociable spotlight as other party for organizing the conversational sequences in talk-in-interaction. In addition, we motivated to explore how the sociable spotlight is embedded within the organization of conversation and how the user’s behaviors are changed according to the sociable spotlight’s behaviors by investigate through conversation analysis of a video-recording. Finally, we conclude how the agency of artificial things constructed in multi-party conversation from minimal designing point of view.

Talking-Ally (Chapter 6) is a system for organizing utterances with hearer’s ‘hearsership’. In this chapter, we mention a social robot that considers the concepts of ‘addressivity’, ‘dialogicality’ and ‘hearsership’ which are considering hearer’s attitude and responses; (1) recognize the eye gaze of the hearer by the eye mark recorder, and (2) state of the hearer to adjust the utterance strategy. As a result, the utterance organized with the hearer is original utterances. In the experiment, we conducted a survey to explore effectiveness of the utterance generation system (social robot). The utterance generated by the proposed approach was able to get

a high commendation for most of the questions. It is suggested that, the utterance of the robot is considered to be equipped with ‘address’. Because the robot was uttered while consciously aware of the eye gaze of the hearer and hearer thinks, “I have been conscious as the hearer”. We got the knowledge that utterance with ‘address’ inspire “intentional stance” from hearer.

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 はじめに	1
1.2 先行研究	2
1.2.1 Human-Robot Interaction(HRI)/ Human-Agent Interaction(HAI)	2
1.2.2 人を揺り動かすことの可能な HRI/HAI の実現に向けて	4
1.3 中心となるアイデア	7
1.3.1 やりとりの中に宿る説得性（説得性の社会的構成）	9
1.3.2 説得性の社会的構成に着目したメディアについて	11
1.4 本研究の進め方と本論文の構成	12
<b>第2章 説得性の社会的構成に着目したメディア技術—タイプ I</b>	<b>15</b>
2.1 はじめに	15
2.2 社会的相互行為における「会話の場」	17
2.3 TableTalkPlus：コンセプトとその背景	19
2.3.1 ヒトとモノをつなぐインタラクションデザイン	19
2.3.2 「場」や「コト」に対するデザインの可能性	19
2.3.3 人と人との共同性を引きだすデザイン	21
2.4 TableTalkPlus の構成と実装	21
2.4.1 ハードウェア	21
2.4.2 ソフトウェア	22
2.5 実験(1)	26

---

2.5.1	目的	26
2.5.2	共同想起対話	26
2.5.3	被験者	26
2.5.4	計画	26
2.5.5	手続き	27
2.5.6	結果	28
2.6	実験(2)	31
2.6.1	目的	31
2.6.2	被験者	31
2.6.3	計画	31
2.6.4	手続き	32
2.6.5	結果	33
2.7	考察	40
2.8	おわりに	42
<b>第3章</b>	<b>説得性の社会的構成に着目したメディア技術—タイプII</b>	<b>45</b>
3.1	まえがき	45
3.2	Optic Flow Field とその社会的相互行為の場への展開	47
3.2.1	視覚的運動制御	47
3.2.2	他者(社会的な環境)との関わりへの展開	48
3.2.3	Optic Flow Field の社会的相互行為への展開	48
3.2.4	TableTalkPlus とその特徴	49
3.3	NEXUS+	52
3.3.1	システム構成	52
3.3.2	ハードウェア	52
3.3.3	ソフトウェア	53
3.3.4	インタラクションデザイン	53
3.4	実験	57



---

3.4.1	目的	57
3.4.2	被験者	57
3.4.3	手続き	57
3.4.4	結果	58
3.4.5	まとめと考察	65
3.5	まとめと今後の展望	67
<b>第4章</b>	<b>説得性の社会的構成に着目したメディア技術—タイプ III</b>	<b>69</b>
4.1	はじめに	69
4.2	The Sociable Spotlights	71
4.2.1	認知的なアーティファクトの内的表象	72
4.2.2	傍観者としての役割を果たすシステムの構成	74
4.3	評価実験	76
4.4	結果	77
4.4.1	記号的表象を介した被験者間のつながり	78
4.5	まとめ	80
<b>第5章</b>	<b>説得性の社会的構成に着目したメディア技術の HAI への応用例</b>	<b>83</b>
5.1	はじめに	83
5.2	実験システムのデザインと構成	85
5.2.1	デザインのポイント	85
5.2.2	システム構成	86
5.2.3	Sociable Spotlight の移動関数	89
5.3	観察 I—公共施設でのインタラクション	90
5.3.1	観察の概要	90
5.3.2	観察手法	91
5.3.3	観察結果	92
5.3.4	まとめ	93
5.4	観察 II—実験室環境における多人数会話	95

---

5.4.1	観察の概要	95
5.4.2	観察手法	95
5.4.3	分析手法	98
5.4.4	観察結果—質的フェーズ	99
5.4.5	観察結果—量的フェーズ	104
5.4.6	観察結果—統合フェーズ	113
5.5	全体総括と今後の課題・展望	115
<b>第6章</b>	<b>説得性の社会的構成に着目したメディア技術のHRIへの応用例</b>	<b>119</b>
6.1	はじめに	119
6.2	研究の背景	121
6.2.1	発話の非流暢性に対する解釈	121
6.2.2	会話における「宛名性」と「聞き手性」	121
6.2.3	設計的な構えと志向的な構え	123
6.3	プロトタイプ的设计と実装	124
6.3.1	Talking-Allyの構築	124
6.3.2	発話のデザイン	126
6.3.3	聞き手の状態のセンシング	126
6.4	動作例	129
6.4.1	インタラクション例1	129
6.4.2	インタラクション例2	130
6.4.3	インタラクション例3	131
6.5	印象評定	133
6.5.1	実験内容	133
6.5.2	結果	136
6.6	インタラクションログによる被験者の行動分析	142
6.6.1	被験者の注視領域の推移	142
6.6.2	ロボットの発話に伴う視線方向の変化	144

---

6.6.3	説得性の効果 . . . . .	146
6.6.4	反応時間の評価 . . . . .	147
6.6.5	被験者—ロボットの相互に与えた影響 . . . . .	149
6.7	総括と今後の展望 . . . . .	151
<b>第7章</b>	<b>結論</b>	<b>155</b>

## 目次

1.1	スウィンググループによる人の揺動運動 . . . . .	7
1.2	光学的流動を再現したアニメーションの一コマ . . . . .	9
2.1	社会的媒介物 . . . . .	20
2.2	ハードウェア構成 . . . . .	22
2.3	クラウド . . . . .	23
2.4	状態遷移と移行適切場 . . . . .	25
2.5	平均発言時間 . . . . .	34
2.6	発言回数 . . . . .	34
2.7	発言割合 (被験者合計) . . . . .	35
3.1	Table Talk Plus の動作例 . . . . .	50
3.2	より多くのクリーチャが自分の方に集まってくる様子を見て発言を 終了する . . . . .	51
3.3	より多くのクリーチャが他者の方に集まっていく様子を見て発言を 躊躇する . . . . .	51
3.4	発言タイミングの調整によりテーブル中央にクリーチャを留めよう とする . . . . .	51
3.5	NEXUS+のシステム構成と内部の処理 . . . . .	53
3.6	NEXUS+の動作例 1 . . . . .	55
3.7	NEXUS+の動作例 2 . . . . .	55
3.8	各パーティのクリーチャ数の時間的な変化と操作区間 . . . . .	56
3.9	各操作区間における発言の平均音量 (tukey 検定) . . . . .	60
3.10	操作区間におけるバックチャネルのカウント数 (パーティ1) . . . . .	62
3.11	操作区間におけるバックチャネルのカウント数 (パーティ2) . . . . .	62
3.12	コレスポネンス分析結果 (パーティ1) . . . . .	64
3.13	コレスポネンス分析結果 (パーティ2) . . . . .	64

4.1	複数人での多人数会話に介在する Sociable Spotlights . . . . .	71
4.2	Sociable Spotlights のシステム構成 (左) と認知的なアーティファクトの内的表象を生み出すメカニズム (右) . . . . .	73
4.3	Sociable Spotlights のシステム構成と認知的なアーティファクトの内的表象を生み出すメカニズム . . . . .	73
4.4	参与役割を交替する目的で会話状態に基づき生成される人工物の振る舞い (図中では behavior(0) → SSB0, behavior(1) →SSB1, behavior(2) → SSB2, behavior(3) → SSB3, and behavior(4) →SSB4 と略記) . . . . .	75
4.5	発話を独占する参加者 (左), 中間的な参加者 (中央), 聞き手に回りがちな参加者 (右) のコレスポネンス分析結果 . . . . .	79
4.6	図中のコレスポネンス分析結果は発話を独占する参加者 (左) と聞き手に回りがちな参加者 (中央) と中間的な参加者 (右) と人工物の内的表象の関係を示している . . . . .	79
5.1	システムの概観 (参加者の取り囲むテーブル上に図 5.2 のアニメーションを投影) . . . . .	87
5.2	環境音の変化のタイミングに合わせて随伴的に揺れ動く Sociable Spotlight アニメーションを映したテーブル . . . . .	88
5.3	左: 手づかみを行うグループ (祖父—孫, 他 4 組), 中央: 引っ張り合うグループ (幼児—幼児), 右: シルエットを利用するグループ (父親—娘, 他 1 組) . . . . .	91
5.4	被験者 A のコレスポネンス分析結果 . . . . .	107
5.5	被験者 B のコレスポネンス分析結果 . . . . .	108
5.6	被験者 C (右) のコレスポネンス分析結果 . . . . .	108
5.7	Sociable Spotlight の動作と沈黙区間における話者選択 . . . . .	111
5.8	Sociable Spotlight の動作に対する被験者の発話行動の変化 (グループごとに算出) . . . . .	114

6.1 Talking-Ally とのインタラクション	124
6.2 Web 上のニュース記事を読み上げる Talking-Ally のシステム構成	125
6.3 3つのサーボモータと iPod touch で構成される Talking-Ally の身体	125
6.4 仮想平面とユーザの視線の推移	127
6.5 評価値の平均と標準偏差 (Q1-Q5)	141
6.6 評価値の平均と標準偏差 (Q6-Q10)	142
6.7 会話の進行に伴う被験者の視線の変化をグラフで図示したもの	143
6.8 ロボットの会話の進行に伴う被験者の視線方向のバラつき (被験者 A)	145
6.9 ロボットの会話の進行に伴う被験者の視線方向のバラつき (被験者 B)	146
6.10 視線回復の成功率の推移 (条件 B ⇔ 条件 D 間での比較)	147
6.11 視線回復に要する時間 [秒] の推移 (条件 B ⇔ 条件 D 間での比較)	148
6.12 視線回復に要する時間 [秒] の推移 (会話の前半⇔後半での比較)	149
6.13 被験者の反応時間とロボットの反応時間 (セグメンテーション 1)	150
6.14 被験者の反応時間とロボットの反応時間 (セグメンテーション 2)	151
6.15 被験者の反応時間とロボットの反応時間 (セグメンテーション 3)	152

## 表目次

2.1 質問紙 (4項目, 自由記述)	27
2.2 質問紙 (4項目, 自由記述)	32
2.3 クリーチャの移動速度に対する 35dB を超える被験者の発話のフレー ム数	37
2.4 分散分析表	38
2.5 Ryan の方法による多重比較の結果	38
3.1 コレスポンデンス分析で使したラベル	63

4.1	スポットライトの各動作に対する被験者の参与役割交替の成功率 (図中では、発話を独占する参加者 → Over, 中間的な参加者 → Average, 聞き手に回りがちな参加者 → Less と略記 . . . . .	78
5.1	抽出の手順 . . . . .	93
5.2	抽出の結果 . . . . .	93
5.3	抽出の手順とエージェンシー付与の判断基準 . . . . .	99
5.4	各グループのエージェンシー形成時の会話の流れとその形成速度 . .	103
5.5	コレスポネンス分析で使⽤したラベル . . . . .	104
5.6	Sociable Spotlight の変化に対する被験者ごとの振る舞い . . . . .	106
5.7	被験者の発話行動の分類 . . . . .	114
6.1	インタラクションの生成方法 . . . . .	128
6.2	トランスクリプト中の注釈の中で使⽤したラベルとその内容 . . . .	129
6.3	実験条件 . . . . .	134
6.4	質問紙 . . . . .	135
6.5	各実験条件における評価値の平均 . . . . .	136
6.6	各実験条件における評価値の標準偏差 . . . . .	137
6.7	t検定の結果 (有意差が現れた項目について) . . . . .	138
6.8	Q1 に対する各被験者の評価値 . . . . .	139
6.9	会話の進行に伴う被験者の視線の変化 (条件 A ⇔ 条件 D) . . . . .	143





# 第1章 序論

## 1.1 はじめに

人の身体を揺り動かしたり、その行動を促すためには、その相手を手で押ししたり引いたりするような (1) 物理的な手段によるものと、その相手に何かをお願いしたり、新たな情報を提供して、相手の行動を促すような (2) 情動的な手段がある。これまで人とロボットとのインタラクションやコミュニケーションを議論する際にも、この二つの手段を中心に議論されてきた [113, 117]。しかし、これまでのロボットが発する「ありがとう」という合成音声にお礼の気持ちを感じることはなく、ロボットからの「助けてー！」という言葉は私たちを揺り動かすような力 (=説得性<sup>1</sup>) を備えるものになっていない [110, 112, 116]。このことは一般のメディア技術においても同様であり、この説得性を備えたメディア技術 (Persuasive Technology) に関しては、最近になってようやく議論が開始された [35] ところである。

本研究は、こうした説得性を備えたメディア技術に向けた研究の一環として、上記の (1) 物理的な手段とも (2) 情動的な手段とも異なる、私たちの身体を揺り動かすことの可能な「説得性の社会的構成」に着目した新たなメディア技術の開拓を目的としている。また、本技術を使用した人とロボット/ソフトウェアエージェントとのインタラクションにおける応用例を示すことも、本研究の目的を実現する上で非常に重要な部分を占める。

本研究の特色は次の二つにまとめられる。一つは、従来の情報処理的アプローチ (コードモデルとしてのコミュニケーション観) とは異なるオリジナルな視点

---

<sup>1</sup>ここでの説得性は“ある個人の態度や行動が変化するように、意図的にデザインされた仕掛け”である [22]。

である。説得性は言葉などのメッセージ（実体）に備わるのではなく、むしろ、その場に参加する参与者同士のやりとりの中で構成されるという視点である。もう一つは、上記の社会的に構成される説得性を、ロボットやインタラクティブメディアを用いて構成論的に生み出しながら、その背後にある原理を探ろうとするアプローチである。専門的なスキルを必要とする社会学や会話分析研究における参与観察・フィールドワークによる手法と比較して、極めてシンプルな形でその性質を明らかにできる点も本研究の特徴といえる [115]。後者の観点は人とロボット、人とソフトウェアエージェントとのインタラクションを扱う Human-Robot Interaction（以下、HRI と略記）研究や Human-Agent Interaction（以下、HAI と略記）の研究分野と関わりが深い。

## 1.2 先行研究

本研究は人とロボット、人とソフトウェアエージェントとのインタラクションを扱う HRI/HAI 研究の分野に位置づけられる。近年この分野は、発達心理学や認知科学、社会心理学などの分野と交流しながら学際的な領域へと進展しつつある [99]。

### 1.2.1 Human-Robot Interaction(HRI)/ Human-Agent Interaction(HAI)

HRIでは、大きく分けて二つの課題にアプローチしている。一つは (I) 産業用ロボットのマニピュレーションや誘導といった制御理論であり、人間の作業者との円滑な協調作業を行うロボティクスに欠かせない基礎技術を議論している [13, 19, 25, 59]。もう一つは (II) 人型ロボットや動物を模したクリーチャなどのソーシャルロボティクス [5] に対するものであり、人とロボットのより親和的・協調的な関係性の構築に向けた課題を議論している。そこではロボットを用いた人の認知/心理実験 [3, 56, 1, 34]、フィールドワークなどの社会実験 [40, 87]、会話研究 [57, 46, 2, 82]

が主に行われている。

とくに、課題IIは後述のHAIにおける中心的な課題とも重なり、本研究が取り扱う問題を内包している。ここでHRIとHAIの違いは次のとおりである。HRIでは物理的な身体を持ったロボットのみを理想的なエージェントとして想定し、その有効性を次のように解釈する。物理的な身体（実体）を有するロボットは人間と同じ空間に物理的に存在できるという点で有利 [39, 66, 74] である。一方、HAIでは、ある協調タスクにおいてロボットが最も有効なのか、他のソフトウェアエージェントやデバイスでは達成できないのかということ自体が議論になる [101]。本研究ではこの点に配慮し、新たなコミュニケーション技術を実装したロボット/ソフトウェアエージェントの両方を用いた応用・展開を示すことで、二つのインタラクションの差異を横断的に検討する。

HRI/HAI研究のもう一つの興味は、新たなコミュニケーション技術を用いたロボット/ソフトウェアエージェントがユーザの行動にどのような効果を与えるのかである。また、これらの知見から人間の行動特性を把握しようとしている。

HRI/HAI研究が一般のユーザや消費者に対して与える寄与は大きく分けて四つある [67]。家庭内や屋内/屋外での物理的 (physical) なアシスト、知覚的 (perceptual) なサポート、認知的 (cognitive) な支援、社会的な関係形成 (social) の援助である。これらは一般的に Assistive Robotics と呼ばれている [18]。そして、この Assistive Robotics は物理的な手段を用いて行われるアプローチ (物理的な手段による HRI/HAI) と情報的な手段を用いて行われるアプローチ (情報的な手段による HRI/HAI) の二つに大別することができる。

### 物理的な手段による HRI/HAI

物理的な手段による HRI/HAI の代表例として、患者の抱き起こしや移動の作業をサポートし、介護者の負担を軽減することを目的としたパワーアシストスーツ [90] や、人のより自然な操作での移動を可能にする全方向移動車椅子 [83] が挙げられる。ベッドまわりの移乗動作を介助するロボットの現状と課題については橋野が詳しく報告している [91]。また、米田は電動車椅子や歩行介助装置について、

現状と課題を明らかにしている [123]。このように、物理的な手段による HRI/HAI では、力仕事のロボットによる代替として、ハードウェアの面から人をサポートする。

### 情報的な手段による HRI/HAI

情報的な手段による HRI/HAI の代表例として、来場者や観客に対して絵画の説明を行うミュージアムガイドロボット [75, 122] や人間を目的地まで連れて行く道案内ロボット [118] が挙げられる。情報的な手段による HRI/HAI では、認知・知覚的な面から、ユーザにとって有益な情報をロボットが提供する。

#### 1.2.2 人を揺り動かすことの可能な HRI/HAI の実現に向けて

ここで再び 1.1 の議論に戻る。相手の身体を特定の場所まで揺り動かしたい。このとき、私たちは一般的に二つの手段を用いる。一つは (1) 物理的な手段であり、相手の身体を手で押したり、引っ張ることで、相手の身体を特定の場所に移動することができる。この手段 1 については、1.2.1 の物理的な手段として、HRI/HAI においても実現可能であることを示した。もう一つは (2) 情報的な手段であり、相手に口頭でお願いをしたり、相手にとって興味深い情報を与えることで、私たちは相手の身体を特定の場所に移動させることができる。この手段 2 についても、1.2.1 の情報的な手段として、HRI/HAI においても実現可能であることを示した。

しかし、私たちの日常的な場面では、手段 1 や手段 2 によらないもう一つの手段が思いがけず用いられることがある。幼児のおぼつかない足どりを見て、私たちはハラハラとしてしまう。その幼児が今にも倒れそうになれば、思わず手を差し伸べて、助けようとしてしまう。よたよたと歩くその幼児に対して、私たちの身体は勝手に動いてしまう [95]。このように、幼児のおぼつかない振る舞いは、他者の身体を揺り動かす一つの方略として捉えることができ、私たちの日常生活の中で観察することができる。一方、HRI/HAI の領域では十分に議論されておらず、未解決の課題といえる。このことから、HRI/HAI におけるコミュニケーションと人々

が日常生活の中で行うコミュニケーションとの間にギャップが生まれ、HRI/HAIにおけるコミュニケーションに現実味が感じられないことがある。このコミュニケーションとしてのリアリティの問題がHRI/HAIの次の大きな課題といえる。ロボットの振る舞いに他者を揺り動かすような力は備わるかという問題である。

ここでは、上記の課題にアプローチするため、関連する国内外の研究動向についてまず整理した。そして、これらの先行研究が明らかにしてきたことについて言及し、本研究の取り組むべき課題を明確にする。

### 説得性を備えたメディア技術 (Persuasive Technology)

ロボットの振る舞いに人を揺り動かすような力は備わるのか。人を揺り動かすような説得性を備えたメディア技術 (Persuasive Technology) に関する研究はFoggらによって提唱され、体系的に行われてきた [20, 21, 22]。この Persuasive Technology の主な狙いはコンピュータによる人の福利厚生や教育、環境保全、安全面での強化である。

福利厚生分野では、ダイエットや体操 [10, 36]、社会関係の維持 [55] を勧めたり、禁煙をサポートする研究 [30] がある。教育分野では、児童の学習に対するモチベーションを向上させる技術 [50] がある。環境保全分野では、ユーザのエネルギー消費を抑える研究が住環境 [54] やスクールバス [65] において実現されている。安全面の強化では、道路上に設置された速度抑制装置 [21] が古典的である。

このように Persuasive Technology は人々の日常的な活動を支援し、Quality of Life の向上を目的として、人間の態度や行動を変化させるように意図的にデザインされた技術要素を扱っている。

### 説得性を備えた HRI/HAI (Persuasive Robotics)

この Persuasive Technology の考え方をロボットと人との関わりに展開する研究 (Persuasive Robotics) がある。国外では ACM/IEEE International Conference on Human-Robot Interaction や International Conference on Social Robotics などの

国際会議において、活発な議論が進められてきた。

例えば、Siegelらはロボットの性別に幅を持たせた場合に、男性の被験者は女性のような振る舞いを行うロボットに献金をする傾向が高いことを明らかにしている [76]。エネルギー消費量の抑制に関する Roubroeks らの実験では、コンピュータからのテキストのみによるアドバイス（説得）と、身体動作を伴わないソフトウェアエージェントによるアドバイス（説得）、身体動作を伴うソフトウェアエージェントからのアドバイス（説得）では、ユーザの心理的な反応がそれぞれの条件で異なることを指摘している [72]。ロボットの身体的・音声的な振る舞いによる人々の意思決定に与える影響・効果（＝説得性）については Chidambaram らの仔細な心理実験が代表的である [8]。また、Ham らは、物語を伝えるストーリーテラーとしてのロボットの視線方向やその身体動作が物語の説得性に与える効果について言及している [32]。

国内では中川や大澤らの研究が代表例として挙げられる。中川らはロボットが能動的にタッチを行う場合と受動的にタッチを受けるだけの場合、タッチを一切行わない場合では、能動的なタッチ動作を伴う方がタスク遂行に関する質が量的に向上し、より友好的な印象をロボットから受けることを報告している [58]。また、大澤らは指先に装着するタイプの眼球を模したデバイスによる擬人化エージェント Pygmy に関する研究発表の中で、エージェントとしての存在感という側面からエージェントの説得性について触れていた [62]。

これらの研究成果からも伺えるように、説得性を備えた HRI/HAI ではロボット/ソフトウェアエージェントによってユーザの態度や意思決定・印象評価にポジティブな影響を与える効果が期待できる。そこで、本研究では思わず身体を揺り動かしてしまうような、無意識的に引き出される振る舞いというより身体レベルでの行動を促す「説得性の社会的構成」に着目して、このような Persuasive Technology/Persuasive Robotics を実現し、ロボットの振る舞いに他者の身体を揺り動かすような力は備わるかという問題にアプローチする。次章以降では、人を揺り動かすことの可能な説得性を備えた新しいメディア技術の実現方法とその HRI/HAI に与える可能性について論じる。

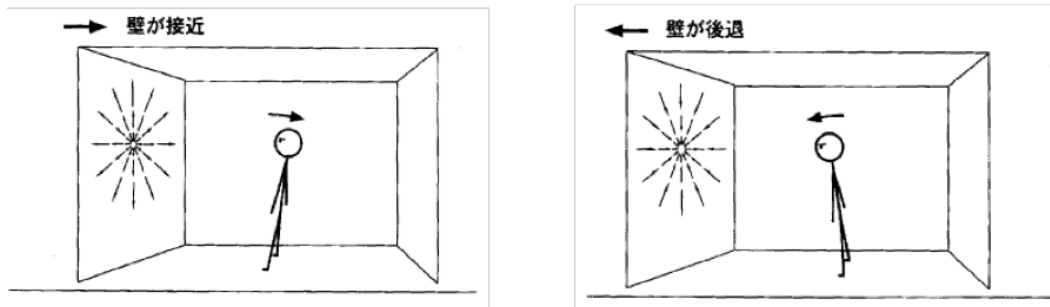


図 1.1: スウィングングルームによる人の揺動運動

### 1.3 中心となるアイデア

これまでの議論のポイントをまず整理し、本研究の中心となるアイデアを述べる。本研究では一つの疑問が研究の出発点になっていた。ロボットからの「ありがとう」という合成音声に感謝の気持ちを感じることはあるのだろうか。ロボットの振る舞いに他者の身体を揺り動かすような力は備わるのかという問題である。

人と人とのコミュニケーションでは、私たちの身体は他者からの語りかけに思いがけず揺り動かされる。一方、人とロボットのコミュニケーションでは、ロボットからの語りかけに私たちの身体が揺り動かされることはない。確かに、ロボットの口からは「ありがとう」という情報は伝わってくる。しかし、そのロボットが本当に感謝をしているのかが私たちと共有できていない。このように、人間同士のコミュニケーションと人とロボットのコミュニケーションとの間にギャップが存在している。ロボットとのコミュニケーションに現実味（リアリティ）が感じられないという問題はこのようなギャップから生じている。そこで、本研究では、ロボットの振る舞いによって他者の身体を揺り動かすという問題にアプローチした。

他者の身体を揺り動かす方法として、これまでに2種類が検討されてきた。一つは、相手の身体を手で押す、引っ張るという物理的な手段である。もう一つは、相手に「どいて下さい」と口頭でお願いする、あるいは、相手に「そこに居ると危ないですよ」と新たな情報を提供して移動を促すような情報的手段がある。この二つの手段は人とロボットとのやり取り（HRI）においても実現可能であった。

そして、相手の身体を揺り動かす手段がもう一つある。図 1.1 は Lee らによって行われた Optic Flow Field に関する研究 [48] のポイントを端的に示した図である。いま、ある被験者が箱形の部屋の中に入り、直立の姿勢で定位<sup>2</sup>している。この部屋（スウィングルームと呼ぶ）は床とは独立して水平方向に移動することができる。ここで図 1.1(左)に示されるように、部屋を右側に移動させる。すると、被験者からは壁が接近しているように見える。このとき、被験者の視界には、内側から外側に向かって視界が移動していくような光学的な流れ（＝光学的流動（optical flow））が現れる。そして、このような内側から外側への光学的流動によって、定位していた被験者は思いがけず後方に倒れかかる。

反対方向の場合についても同様である。図 1.1(右)に示されるように、部屋を左側に移動させる。すると、被験者からは壁が後退していくように見える。このとき、被験者の視界には、外側から内側に向かって視界が移動していくような光学的流動が現れる。このような外側から内側への光学的流動によって、部屋の内部で定位していた被験者は思いがけず前方に倒れかかる。このように、光学的流動の人為的な操作によって、人の揺動運動を誘発させることができる。

この Optic Flow Field に関する研究の最も重要な指摘は、人の行為（調整）が環境と一体になって形づくられているという点である。従って、図 1.2 に示すような光学的流動を再現したビデオ刺激によって人の揺動運動を引き出せるという主張とは異なることに注意したい。

私たちは環境にある情報（例えば、光学的流動の変化）と関係を結び、自分の行為を調整するためのリソース（行為者から参照される情報）として活用している。このリソースとうまく同期をとり、その情報の変化に合わせながら自分の行為を絶えず微調整し続ければ、（微小な逸脱を繰り返しつつも）当初に思い描いた一連の行動を産出することができる。ここで、行為の調整を繰り返している最中に、そのリソースに対して第三者による人為的な操作が加えられた場合では、行為者の身体はその調整の途中であって、別の方向に揺り動かされてしまう。このような観点が本研究の中心となるアイデアであり、このアイデアに従って、人と

---

<sup>2</sup>その場でじっとするという意味。



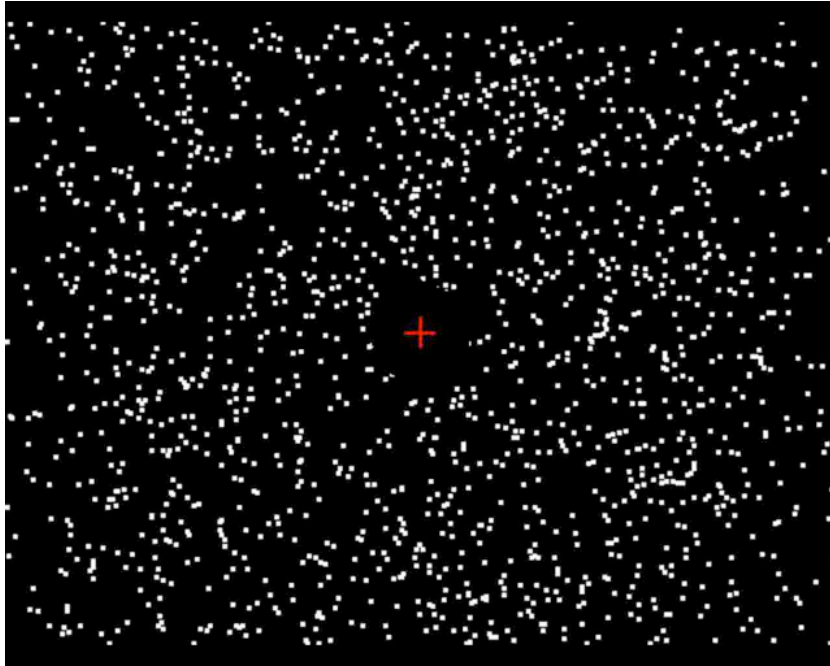


図 1.2: 光学的流動を再現したアニメーションの一コマ

ロボットのインタラクションにおいて、人の身体を揺り動かすことが最終的な目標である。このような手段によって人の身体を揺り動かすことは、HRIの研究分野において未開拓の領域であった。

### 1.3.1 やりとりの中に宿る説得性（説得性の社会的構成）

上述のアイデアは、相手とどのような関係を形成するのかによって、同一の言葉であっても揺り動かされる場合と揺り動かされない場合がある（＝説得力が変わる）ということを示している。これは私たちの日常生活に例えると分かり易い。例えば、コンビニエンスストアでのマニュアル通りの接客では、「ありがとうございました」と言われても、相手の感謝の気持ちを感じとることができない。しかし、相手が自分に対して深く会釈しているような場面で「ありがとうございました」と言われれば、相手の感謝の気持ちを直に感じることができる。同一の「あ

りがとうございました」という言葉であっても、相手とどのような関係を形成した中で繰り出された発話なのかによって説得力が変わっている。

このように、人を揺り動かすような説得性は言葉などのメッセージに宿るのではなく、むしろその場に参加する参与者との関係の中で社会的に構成されるものである。この考え方はウィットゲンシュタイン (L. Wittgenstein) の言語ゲームのメタファから借りている。この言語ゲームのエッセンスはチェスのコマの働きによく例えられる [104]。ここでは説得性というレベルでの考え方を [104] の解釈に加えて、本研究のキー概念である説得性の社会的構成について明確に述べる。

二人の参与者がさまざまな形状・大きさのコマをチェッカーボード上で交互に動かす。これがチェスと呼ばれるゲームである。このとき一つひとつのコマを、いつ、どのように動かすことができるのかという明示的なルールが、明示されていないルール<sup>3</sup>と一緒に混在している。その上で、チェスの一つひとつのコマの動きはゲーム全体から制約を受ける。しかし、木製の小さな複数のコマはチェスというゲームの外では、その動きを制約するものはない。ひとたび、チェスゲームに戻れば、最も小さなコマでさえ、チェックメイトという掛け声とともに、キングやクイーンの動きを抑制することのできる力 (= 説得力) を備える。チェスゲームにおける木製のコマは二人の参与者が企てたゲームの中で用いられることによって、小さなコマの動きが大きなコマの動きを抑制するという説得性が備わる。

ここでの主張は、説得性は実体としてのモノにはじめから属性として備わっているのではなく、参与者同士のやりとりの中で構成されるものなのではないか (= 説得性の社会的構成) という考え方である。この観点に基づき、社会的に構成される説得性がリアルタイムに生み出されるような実験環境を三つデザインした。また、この実験環境をインタラクティブメディアとして提案した。さらに、これらインタラクティブメディアの特性を、新しいコミュニケーション技術としてロボット/ソフトウェアエージェントに適合する形で実装し、HRI/HAI への応用・展開を示した。

<sup>3</sup>相手を罵倒してはいけないという不適切な振る舞いを抑制するルール

### 1.3.2 説得性の社会的構成に着目したメディアについて

上記の三つのインタラクティブメディアの設計やそのHRI/HAIへの展開・応用に際しては、次の点に考慮した。ここでは、上記アイデアの実現手段について整理する。

本技術の実装に際しては、次の四つの要素が重要である。(1) 説得性が社会的に構成されるような状況設定, (2) リアルタイムに説得性が生みだされる仕掛け, (3) 被験者のインタラクションログデータの分析に当たっては身体レベル・行動レベルでの分析 (behavior) に重点をおく, (4) 複数タイプのインタラクティブメディアによる実験結果に対して横断的な検討を行い、新しい知見を得ることである。

要素の1は、インタラクティブメディアのハードウェアに関する要件である。人と人の中で説得性が社会的に構成されるためには、複数人の参与者による何らかのやりとり (=相互行為) を行わなければならないという状況を設定する必要がある。この点から、インタラクティブメディアのハードウェアは複数人が同時に参与できること。また、このインタラクティブメディアを介して、何らかの相互行為を行う必要がある。ここでは、三人での多人数会話を想定し、そこでの相互行為のやりとりを媒介するインタラクティブアートとして実現した。

要素の2は、このインタラクティブアートのデザインに関する要件である。リアルタイムに説得性が生み出されるためには、言語的な要素や意味などの情報的な側面 (コンテンツの部分) を削ぎ落とす必要がある。なぜならば、私たちの社会的な慣習の中で定着している日常的なデザインを用いることは、既に人々の間で社会的に構成されていた説得性の効果を再確認するだけだからである。このことに配慮し、ここでは参与者がそのデザインをどのように解釈すれば良いのか、どのように用いて良いのか分からないような抽象的なデザインを検討した。また、このようなデザイン手法を「引き算としてのデザイン」と呼んでいる。やりとりの中で立ち現れてくる新たな側面を顕在化することができる [78, 111]。

要素の3は、本技術の評価方法に関する要件である。リアルタイムに説得性が社会的に構成される場合、その相互行為に参加している参与者はその説得性の存在を自分の身体によって感じるができるという。ただし、その説得性はそこ

での相互行為の中で構成されたばかりのものであり、参加者の間で定着しているわけではない。従って、その説得性の効力を個人がはっきりと認識し、また、その効力を言語などのコミュニケーション手段を通じて、他者に明確に伝えることができるかは定かではない。そこで、インタラクティブメディアの評価、すなわち、説得性の効果の検証に際しては、被験者のインタラクションログデータを可能な限り行動レベルで記録・観察・分析した。また、このように無意識的に身体が揺り動かされる現象を説明する場合に、「場」という概念を使用した。お互いの行為を制約するような社会的な「力」を備えるものである。

要素の4は、被験者の行動データを収集/比較できるデータ分析の環境を用意することで、それぞれの特徴や限界を調査した。本技術の実装では、以上の4つの要素に配慮している。一方、本研究の推進にあたっては、次の事項に配慮した。

## 1.4 本研究の進め方と本論文の構成

本研究を進めるに当たり、次の二つのステップを設けた。第一のステップでは、上記の観点から、他者を揺り動かすことの可能な説得性を、人と人のやりとりを媒介するインタラクティブメディアを使用して生み出し、その背後にある原理を明らかにする—構成的理解のステップ。二つ目のステップでは、構成的理解のステップにおいて新たに確立したコミュニケーション技術を Human-Robot Interaction(HRI)/Human-Agent Interaction(HAI) に応用・展開する—応用・展開のステップ。

次の三つの手順により、構成的理解のステップを達成する。一つは (I) 理論的な基盤の構築である。そのために、説得性を備えたメディア技術に関する研究動向をまず把握した (1.2 節)。そして、本研究でアプローチする説得性の社会的構成に必要な要素、そのメカニズムについて、理論的基盤を整理した (1.3 節)。二つ目は (II) I の理論的基礎を検証するための研究プラットフォームの構築である。本研究のプラットフォームとして、生態心理学で議論されてきた包囲光配列のアナロジーを会話などの社会的相互行為に展開する Table Talk Plus (2 章)、ネットワークを介して促進行為場を生み出す NEXUS+ (3 章)、Goffman の「傍観者

としての参与」の考えに基づく Sociable Spotlight (4章) を構築した。三つ目に (III) 私たちを揺り動かすことの可能な説得性を備えたコミュニケーション技術を確立した。そのために、上記IIのインタラクティブメディアについて、実験室内における被験者実験や近隣の交流施設における観察を通じ、人の行動を揺り動かすような説得性の効果を確認するとともに、人と人のインタラクティブメディアを介した説得性を伴う新しいコミュニケーション技術として提案する。

応用・展開のステップでは、HRI/HAIにおける応用例を示す。上記IIIにおいて確立した人を揺り動かすことの可能な説得性を有するコミュニケーション技術を、実際のロボット/ソフトウェアエージェントに適合する形で実装する。ここではHAIにおける一つの応用例を示す。具体的には、認知科学の分野で議論されてきた Heider&Simmel Demonstration の原理を、人と人との社会的な関わりに展開する Sociable Spotlight アニメーション (ソフトウェアエージェント) を構築した (5章)。二つ目にHRIにおける一つの応用例を示す。ここでは聞き手 (ユーザ) との相互行為的調整に基づく発話生成エージェント Talking-Ally を構築し、人とロボットの説得性を伴ったコミュニケーションを実現した (6章)。最後に、本研究で扱ったインタラクティブメディア/ロボットの特性を横断的に分析し、本研究が明らかにしたことと今後の課題を明確にする (7章)。

次章以降では、これらの事柄について、インタラクティブメディアとしての提案 (2章, 3章, 4章), HAIへの応用・展開 (5章), HRIへの応用・展開 (6章), 横断的な検討・考察 (7章) の順で説明する。



## 第2章 説得性の社会的構成に着目したメディア技術—タイプI

### 2.1 はじめに

これまでコミュニケーション研究の多くは、情報をいかに正確に効率よく伝えるか、メッセージの意味をどのように解釈するか、相手をどのように説得すべきか、といった考え方で進められてきた。

その一方で、私たちの日々の雑談などを考えてみると、その目的は必ずしも正確な情報の伝達に限られない。むしろ、言葉のやり取りそのものを楽しんだり、一緒にそうした場や話題を共有することそのものを志向しているともいえるだろう。あるいは、そうした場を共有することを介して、お互いの社会的なつながりを維持し、調整しているものと考えられる。

発達心理学者の鯨岡は、他者との関わりを維持しつつ、同時に自己の欲求の実現を図ろうとする二面性を「自己充実欲求」と「繋合希求性」という二つの言葉で説明している [95]。同様に、言葉や会話もまた、何かを相手に伝え、自分の行為の目的を遂行しようとする側面と、同時に他者とのつながりを維持しようという側面を備えているようである。

Lotman は、コミュニケーションには、相手に何かを伝達するという単声的な要素と、他者との相互行為の中での意味生成的で、多声的な要素とが同居していると指摘している [89]。上司から部下への指示などは、単声的な要素が優位となっている例だろう。一方、日々の何気ない会話では、むしろ他者の言葉の重なりの中でたち現れる新たな意味を楽しむような、意味生成的な側面が優位となっているようである。そこでは、一方的な伝達ではなく、むしろ繋合希求的な側面である

他者の言葉への呼応やそのつながりを重視していると考えられる。

また、Malinowski のいうファティック・コミュニオンは、やり取りすることそのものを目的とするような振舞いの存在を指摘したものである [53]。年賀状などのやり取りは、その伝達内容が問題なのではなく、むしろやり取りすることそのものに意味のあるような、他者との社会的つながり (social bond) を志向したコミュニケーション行動といえるものだろう。

社会言語学者の Tannen (2001) は、こうした会話の性質を男女間の会話スタイルの違いに着目し、レポート・トーク (情報交換に主眼を置いた会話のスタイル) とラポート・トーク (心的なつながりを重視する会話のスタイル) という言葉で説明している [80]。

例えば、「今日は、どこに行ってたの?」、「友達の家」といったお父さんと娘との会話は、レポート・トークと呼ぶものである。一方、お母さんと娘との会話にある、「今日の気分は何色?」、「うーん、今日は曇り空かなあ」といったやり取りは、ラポート・トークであり、何かを共有し合って、つながりを確認することに価値を見出すものである。

筆者らは、人と人との間を媒介し、社会的相互行為の「場」における共同性や社会的なつながりを引き出すような社会的媒介物 (social mediator) について議論している。例えば、母親と子どもとの間にある絵本は、お互いの注意や話題を制約しつつ、コミュニケーションを促進する上で有用な媒介物であろう。実際、関係性発達障害児と養育者との間にある玩具やロボットなどを伴った場は有用な媒介物として機能することが知られている。

本研究では、日常での会話において参加者の間に社会的媒介物として介在し、情報の伝達を促進するのではなく、むしろその参加者間の共同性や社会的なつながりを引き出すようなインタフェースについて検討している。

その一環として、ここでは多人数会話に見られる発話交替などのダイナミクスの可視化を試みる TableTalkPlus と呼ぶシステムを構想し、その実装を進めてきた。会話の中でやり取りされる発話の背後で「地」としてあった、会話のダイナミクスをあえて「図」として可視化することで、会話における共同性や参加者間



でのつながりの要素をデフォルメ（誇張）することを狙うものである。本研究では、その第一段階として、会話の中での TableTalkPlus の存在が会話参加者の発話行動にどのような影響を与えるかを明らかにすることを目的としている。

本章では、はじめに会話における参加者同士の関係性やそこからたち現われるダイナミクス（お互いの行為を制約するような社会的な「力」を備えるもので、これを本論文では「会話の場」と呼ぶ）について、Garfinkel や Goodwin, Goffman などの社会的相互行為論に依拠しながら、その理論的な背景を整理する（2.2 節）。つぎに TableTalkPlus のコンセプトや背景をインタラクションデザインとの関連で整理する（2.3 節）とともに、具体的なハードウェア、ソフトウェアの実装内容について述べる（2.4 節）。また、2つの実験（2.5 節、2.6 節）によって TableTalkPlus の基本動作の確認と TableTalkPlus の存在が会話参加者の発話行動にどのような影響を与えるかについて検証を行った。最後に考察（2.7 節）、とまとめ（2.8 節）を述べる。

## 2.2 社会的相互行為における「会話の場」

社会学者の Garfinkel らによって 1960 年代から進められてきたエスノメソドロジーや会話分析といった研究では、現実の会話現象を説明する様々な概念や装置（モデル）がこれまでに提案されている。これらは、社会的相互行為としての会話（Talk-in-interaction）に焦点が当てられてきた。例えば、Sacks の指摘した「日常会話の順序どりシステム（the turn-taking system for conversation）」は、会話研究の中では最も頻繁に参照される会話モデル（装置）である [85]。私たちの日常的な会話では、大抵は「一時に一人が、そして一人だけが」話している。いつ誰が話すかあらかじめ決まっているわけでもないのに、私たちはこれを当然のこととして受け入れている。Sacks によれば、私たちは社会的な相互行為に参加しながら、お互いの順番の交代を相互に共同して調整することで、そのつど局所的に、会話の参加者自身によって、相互行為的に会話をデザインし、成立させているというのである。

さらに、Goodwin (1979) は、会話における「聞き手」の重要性を次の2点にまとめ、会話が相互行為的なものであることを示した [28]。すなわち、(1) 話し手であることは聞き手の注目に支えられていること、(2) 話し手は聞き手に合わせて、その聞き手が聞いてくれるように「そのつど」自分の発話をデザインしなくてはならないことである。つまり、私たちがいま「話し手」であるのは、相手が「聞き手」になってくれているという単純な事実によって支えられている。私たちは、発話を繰り出すという行為によって「話し手」になれるわけでも、また話を聞いているだけで「聞き手」になれるわけでもなく、その役割までも相互行為的に組織されるということである。

本論文における「会話の場」とは、この「相互行為的かつ拮抗しあう参与者間の関係性やダイナミクス」を指すこととする。例えば、私たちは知り合いに挨拶されたときに思わず挨拶を返してしまうように、相手からの発話に対する何らかの応答責任を無意識に感じてしまう。これは、そのときの「会話の場」によって引き出された共同性の存在によるものだろう [114]。そうした「会話の場」は Goffman のいう「参与フレーム (participation frame) におけるダイナミクス」や、Reed のいう「促進行為場 (promoted action field[68])」の概念に近いといえる。このような「会話の場」は実体としての「モノ」ではなく「コト」であり、私たちの目で直接その存在を確認することは難しい。本研究で提案する TableTalkPlus はこの「会話の場」を支える情報やダイナミクスを、近似的あるいは擬似的に視覚化することを目指している。

また、私たちの行為を方向付け、制約するようなアーティファクトとしての役割を顕在化できないかと考えている。例えば、母親と子どもとの間にある絵本は、二人の会話が向かう対象であると同時に、二人の会話の流れを方向づけ、制約するものである。同様に、私たちの繰り出す発話は私たちの思考が向う対象であり、私たちの思考を制約し、方向づけるアーティファクトとして機能している。「会話の場」を支える情報やダイナミクスを視覚化したり、人為的に変形を加えることで、「会話の場」そのものを新たにデザインすることが可能になると思われる。それにより、会話における伝達的な側面ではなく、むしろ他者との共同性や社会的な

つながりをより促進するような場へと変容させることが可能となると考えられる。

## 2.3 TableTalkPlus : コンセプトとその背景

### 2.3.1 ヒトとモノをつなぐインタラクションデザイン

近年、インタラクションデザインの開発手法に焦点を当てたプロダクト、コンテンツの開発が非常に盛んである。日々新たな製品が生み出され、製品とその利用経験が複雑化していく時代背景の中で、インタラクションデザインに注目が集まることは自然の成り行きだろう。

インタラクションデザインの適用分野は多岐に渡るが、中でもデジタルコンテンツやインタラクティブアートの世界では、表現とユーザインタフェースが直結しているため、ユーザとのインタラクションをデザインする必要性が必然的に高まる。そこでフィジカルコンピューティング (Physical Computing) という考え方が10年ほど前から提案され [23]、注目を集めている。これは従来のマウスやキーボードによる入力よりも、より直感的で体感的な「モノ」を求めるもので、様々なセンサを用いて人間の自然な振る舞いでコンピュータとコミュニケーションすることを目指している。

### 2.3.2 「場」や「コト」に対するデザインの可能性

インタラクションデザインにおける次のターゲットは「モノ」に関する議論から、むしろ人と人との間にある「場」で生じる「コト」のデザインに向けられていくと考えている。2.2で述べたような「会話の場」における応答責任はなぜ生じるのかという問題は、言葉の内容などの「モノ」ではなく、その「場」で生じている「コト」に関するものであろう [114]。

本研究で提案する TableTalkPlus は、人と人が会話を行っているところに介入しながら、そこでの「会話の場」にゆるやかに影響を与えることを狙うものであり、その意味で「コト」をデザインすることを目指している。具体的には、会話

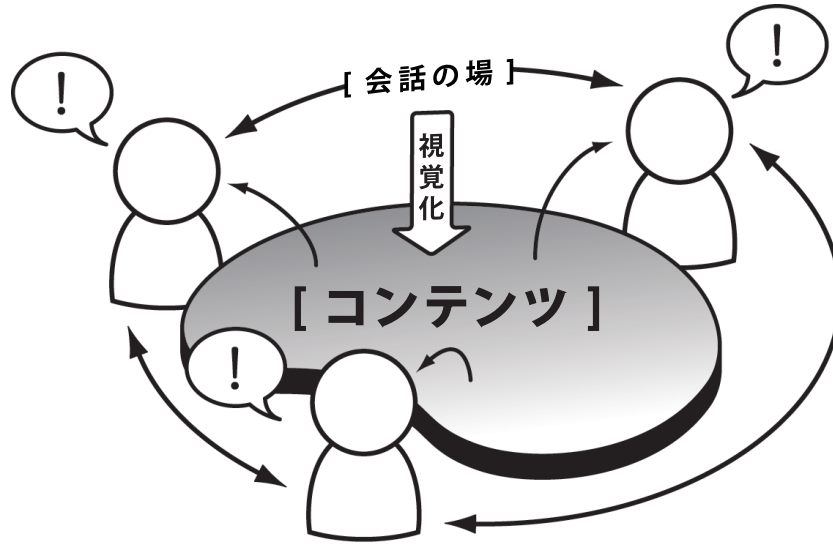


図 2.1: 社会的媒介物

の中での発話の一部の情報を取り込んで、それをCGやアニメーションの手法を用いて、会話参加者の取り囲むテーブルの上にリアルタイムに投影することとした。このような形態を取ることで、次のようなことが可能となると考えられる。

(1) 日常的なオブジェクトを介した、比較的単純な仕組みによって、会話参加者にゆるやかに影響を与える。(2) テーブルは参加者らにとって共有物であり、そこにCGを映し出すことで改めて「場」への共同参加の意識をもたらす。(3) リアルタイムに生成されるCGを見ながら会話することで、それらコンテンツから何らかのフィードバックを受ける。このときコンテンツは進行中の会話と相互に制約しあうアーティファクト (artifact[105]) として機能する。(4) 発話における一部の特徴量を選択し、CGに反映させることで、会話における一部の側面をデフォルト化する。次の章で説明するように、TableTalkPlusは、発話音声における継続時間などの情報を用いることから、言語的な情報の背後にあったタイミングや共同性、会話における継続性を志向する側面などが顕在化されると考える。(5) 「会話の場」をCGで表現する際に、動きのスピードなど、人為的に変更を加えることで、その「会話の場」に与える影響などを統制する。

特に(3)は重要で、コンテンツをアーティファクトとしてとらえることで、会話

そのものを方向づける,あるいは会話にリズムを与え活性化させる,といったポジティブな影響を与えることも可能となると考えられる。このように,会話参加者の間に介在することで,会話の背後にある共同性や社会性をより明示的に示しコミュニケーションを促進させるなどの働きを備える社会的媒介物(図 2.1 Social Mediator)として機能することが期待される。

### 2.3.3 人と人との共同性を引き出すデザイン

本研究の TableTalkPlus と関連する研究として, Ishii らによる “PingPongPlus” [38] がある。PingPongPlus は, 卓球台にピンポン玉の落下位置と連動したグラフィックスを投影するというシステムである。卓球という実世界でのゲームと, そのゲームの流れに呼応して卓球台上に現れるグラフィックとを重ね合わせることで, もともと競技であったはずの卓球が, いつしか卓球のラリーを継続させて楽しむためのゲームに変えることを狙いとしている。PingPongPlus が, 卓球そのものがもつ競技としての側面を変容させ, むしろ共同行為としての側面を顕在化させることで, 参加者の共同性を引き出す社会的媒介物になると考えられる。

会話もまた参加者間で互いに調節しあうゲームのようなものであると捉えれば, PingPongPlus が卓球を介在させたのと同様に, TableTalkPlus を人と人とのコミュニケーションに介在させることで, 会話という本来は情報伝達的な行為であったものを, 「やりとりすることそのもの」に意味があるという共同行為としての側面を顕在化させることで, 会話の性質そのものを変容させることができるのではないかと考えている。

## 2.4 TableTalkPlus の構成と実装

### 2.4.1 ハードウェア

TableTalkPlus は 3 人での多人数自由会話を想定したシステムである。システムを構成するハードウェアの概略図を図 2.2 に示す。

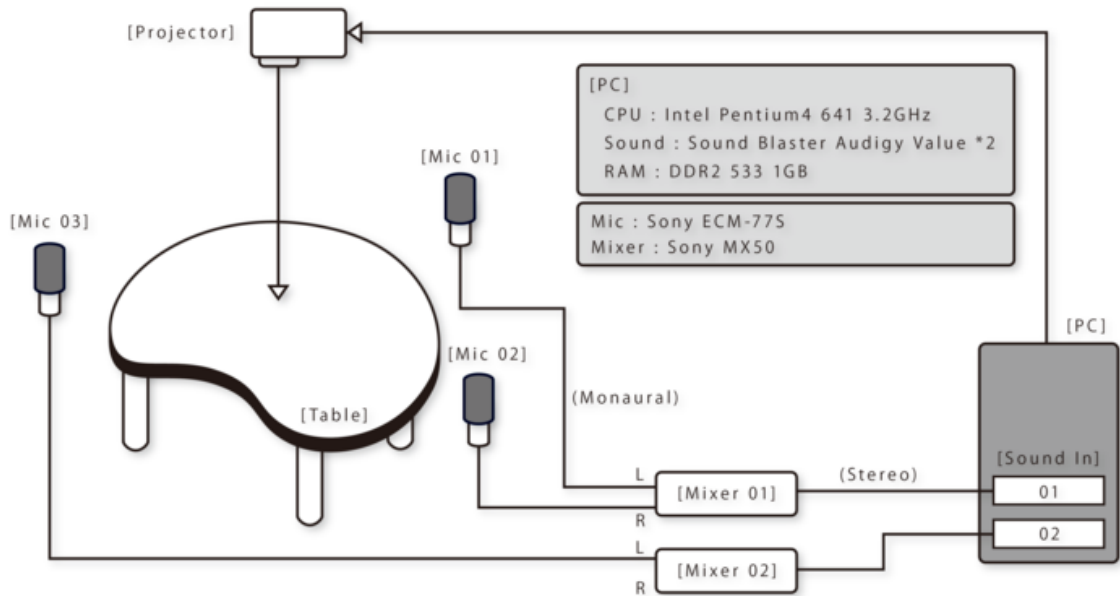


図 2.2: ハードウェア構成

TableTalkPlus のハードウェアは音声入力用マイクが3本、マイクをPCに接続するためのミキサーが2機、制御用PC、映像出力用プロジェクタと映像の投影対象となるテーブルからなる。

### 2.4.2 ソフトウェア

2.2節で述べたような「会話の場」のダイナミクスを表現するため、TableTalkPlus上で動作するアプリケーション（Crowdと呼ぶ）を用意し、「会話の場」を支える情報やダイナミクスの視覚化を試みる。

Crowd（図 2.3）は生命感をイメージしたCGによって、会話における拮抗した関係性やダイナミクスの視覚化を狙う。テーブルには多数のクリーチャが小さな白い三角形のグラフィックとして映し出される。このクリーチャたちは生命感の表現のためにデザインされたもので、人工生命研究の1つである Reynolds による Boids[71] を参考にしている。クリーチャの形が三角形であるのは、それらのアド

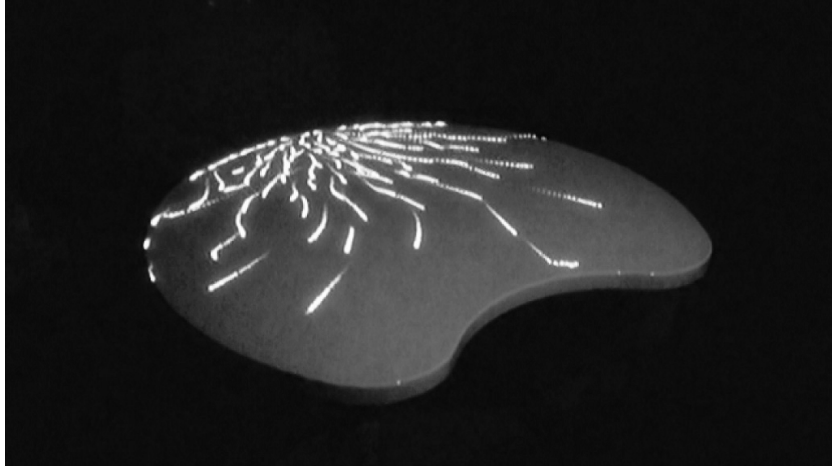


図 2.3: クラウド

レスや進行方向を示すことができる最もシンプルなデザインだからである。

## Crowd

Crowd のクリーチャは次の 2 つのルールに従って動いている。

1. 近くの個体からなるべく離れる
2. 音声入力のある一番近いマイクの方へ向かう

これらのルールは単純なものであるが、局所的な個体同士の相互作用の結果、集団として複雑な動きを見せるものとする。この動きにより参加者同士の会話という相互行為によって動的に変化する「場」を表現する。

## アルゴリズム

Crowd のアルゴリズムについて述べる。  $n$  回目の描画時点での各個体の速度ベクトルの方向  $\theta_n$  は (2.1) 式によって決定される。

$$\theta_n = \theta_{n-1} + C_g \phi_{gn-1} + C_a \phi_{an-1} \quad (2.1)$$

ここで  $\phi_{gn-1}$  はその個体からみた最も近い発話者への相対的な方向（誰も発話していないときは  $\phi_{gn-1} = 0$ ）， $C_g, C_a$  はそれぞれ定数である。 $\phi_{an-1}$  は前回の描画時における個体の速度ベクトルと，(2.2)式で決定されるベクトル  $\vec{a}$  とのなす角である。

$$\vec{a} = \sum_{i=1}^N \frac{\vec{p} - \vec{p}_i}{|\vec{p} - \vec{p}_i|^2} \quad (2.2)$$

ここで  $\vec{p}$  は個体の座標， $\vec{p}_i$  は近くの  $N$  個体の各座標である。 $\phi_a$  によって，クリーチャは近くの個体からなるべく離れる方向に進み，集団として複雑な動きを見せるものとする。

### 状態遷移と移行適切場

Crowd では，参加者の発話に合わせてクリーチャの状態を次のように設計，遷移させている（図 2.4）。

**ON** 参加者が発話している（マイクに閾値以上の入力がある）状態。発話者のもとにクリーチャが集まる。

**OFF** 誰も発話していない状態。テーブル上をバラバラに散らばりながらさまよう。

**OFFHOLD** 発話が終了した直後。発話者の近くに集まったクリーチャはややとどまり，ゆっくりと散らばる。

2.2 で挙げた Goodwin の例で，会話における「聞き手」の重要性を示した。話し手は聞き手の注目に支えられている。これを意識して，上記の設計における ON 状態がデザインされた。ON 状態でクリーチャは発話者の近くに集まる。これが話し手にとっては，周囲の人が自分の話に注意を向けてくれている，という一種の「聞き手性 (hearership)」の表示につながるものとする。

また Sacks の例で，私たちは発話権の交代，すなわち発話の終了をかなりの精度で予測できることを述べた。この時，発話者が一旦発話を終了すると，(会話内容の



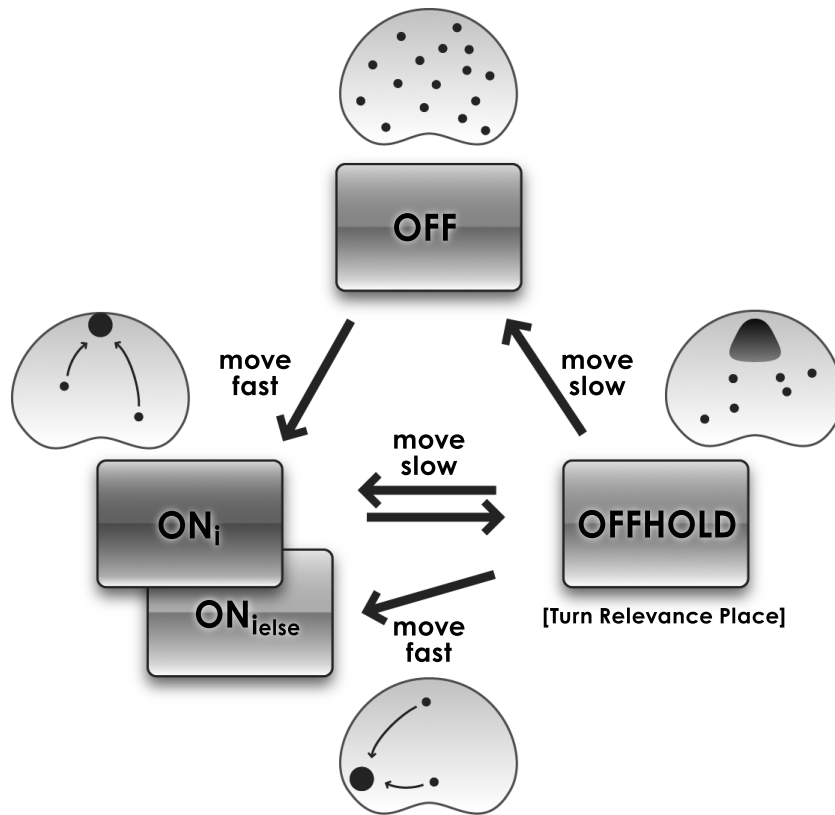


図 2.4: 状態遷移と移行適切場

コンテキストにもよるものの)「次に誰が喋ってもよい」状態が訪れる。このような状態は移行適切場と呼ばれており、これを意識して上記の設計における OFFHOLD をデザインした。具体的には、ON 状態で集まったクリーチャが OFFHOLD でゆっくりと離れていくことによって移行適切場を表す。それにより、参加者に思わず応答責任を感じさせてしまうような、次の発話を喚起させる「場」の表現を狙う。

OFFHOLD 状態に移行して数秒後、クリーチャは OFF 状態に移る。このとき、集まっていたクリーチャがバラバラにさまよいながらテーブル全体に散らばる。この不定な様子によって、誰でも自由に、気軽に会話に参加できる「場」の表現を狙う。

## 2.5 実験(1)

### 2.5.1 目的

本実験では、実際の会話に TableTalkPlus を介在させて、会話参加者が TableTalkPlus をどのようなものとして捉えたか、また各クリーチャの動きを会話の中でどのように意識したかを質問紙への回答に基づいて明らかにする。

### 2.5.2 共同想起対話

本実験では課題として、視聴したビデオに対する共同想起 (joint remembering[16]) 対話を用いた。この方法は各被験者の発話の自由度を保ちながら、自由会話や一定のテーマによる議論などの会話課題に比して会話の内容をある程度統制することができる利点を持つ。

### 2.5.3 被験者

男性 35 名、女性 1 名、計 36 名 (20–25 歳、平均 22.9 歳) の学生が実験に参加した。被験者は相互に顔見知りである 3 人一組の会話集団 (全 12 組) に組み込まれた。なお、全ての被験者は TableTalkPlus を使用しての会話は初めてであった。

### 2.5.4 計画

2.3 で述べた Crowd を使用し、クリーチャが発話に対して反応する度合いを条件とした。各クリーチャの進行方向を決定する (2.1) 式中の係数  $C_g$  によって、反応の度合いが決まる。本実験では、次に示す 3 つの条件を用意した。

**条件 A**  $C_g = 0$  (反応なし)

**条件 B**  $C_g = 0.2$  (弱い反応)

表 2.1: 質問紙 (4項目, 自由記述)

質問番号	質問内容
1	各セッションで何か違いを感じましたか？
2	光が自分のところに集まってきたときに何か感じましたか？
3	光が自分のところから離れていくときに何か感じましたか？
4	その他, 実験について気になったことや感想など

### 条件 C $C_g = 0.8$ (強い反応)

1人の被験者が全ての条件について実験を行う被験者内計画とし, 実験集団間でカウンターバランスをとって順序効果の相殺を行った。

#### 2.5.5 手続き

被験者は実験開始前に, 「ビデオの内容を覚えてストーリーを正確に話す」という課題を伝えられた。実験は1回の練習セッションと, 各実験条件に対応する3回のセッションからなる。各セッションにおいて, 被験者はまずビデオ視聴ブースで5分 (練習セッションは3分) 程度のビデオを見せられ, 次に TableTalkPlus が設置された会話ブースに移りテーブルの周りに着席, 各自が TableTalkPlus に接続されているタイピン型マイクを装着した後, 実験者の合図によって5分 (練習セッションは3分) の間, 先ほど観せられたビデオの内容を思い出しながらストーリーを話し合った。会話中は TableTalkPlus の実験プログラムによる映像をテーブル上に投影した。5分間の会話が終わると, 次のセッションに移った。全てのセッションが終了した後, 質問紙 (表 2.1) に回答し, 実験者によるデブリーフィングを行って実験を終了した。

### 2.5.6 結果

質問紙の各質問項目ごとに多く見られた回答や特徴的な回答を示し、簡単な考察とあわせて以下に述べる。

**質問1**：各セッションで何か違いを感じましたか？

セッションごとの TableTalkPlus のクリーチャの動き方の違いについて記述した次のような回答が多数（23名）みられた。

- 「最後のセッション<sup>1</sup>では光が動きませんでしたね」
- 「光の集まり方がセッションごとに違った。」
- 「1回ごとに、光が自分のところに集まってくる感度が違った気がした。実験の各セッションごとに感度がよくなっていたと思う。」
- 「2回目のセッション<sup>2</sup>では、光が集まらないと感じた。」
- 「下の光の動きが違った気がした。」

以上のように本実験で用意した条件の違いを多くの被験者は認識できている。また、クリーチャの動き方の違いを感じた被験者のうち、半数近く（11名）の被験者はA条件のとき（クリーチャの反応がないとき）に、話しづらさを示すような、「最後のセッション<sup>3</sup>で妙に話しづらかった」という違和感を訴える回答をすることが質問紙から明らかとなった。ここから TableTalkPlus のクリーチャが被験者らの発話に無反応に動くとき、実際の発話に何らかの悪い影響を与えているものと考えられる。

逆に、「光が声に反応することで会話が盛り上がった気がする」など、クリーチャの動きが、会話の活性化につながったことを示す回答があった。これらの回答から、TableTalkPlus のクリーチャのふるまいが会話の性質を変容させるための有用な要素であると考えられる。

<sup>1</sup>このグループの最後のセッションはA条件（反応なし）であった。

<sup>2</sup>このグループの2回目のセッションはA条件（反応なし）であった。

<sup>3</sup>このグループの最後のセッションはA条件（反応なし）であった。

**質問 2：**光が自分のところに集まってきたときに何か感じましたか？

次のような、自分が話し手となっていることを自覚したり、他者の注意を感じるなどの回答が多数（22名）得られた。

- 「自分が話していることを実感した。」
- 「自分がしゃべっているという実感を持てた。」
- 「自分に注目が集まっている感じがして、話しやすかった。」
- 「自分が発言をしているんだなあということが明確にわかった。」
- 「自分がほかの人より目立っていると感じた。」

これらの回答から、TableTalkPlus のクリーチャが ON 状態のときに現れる流動（発話者の近くに集まる）が、被験者らに他の被験者が自分の話に注意を向けてくれている、という一種の「聞き手性」の表示をアシストしたものと考えられる。

また、次のように話し手の交替につながる感情を示す回答も得られている。

- 「早く話し手を交替しなくては、と思った」
- 「少々、自分がしゃべりすぎたかと（思った）」

さらに、次のようにクリーチャを集めることに対するポジティブな感情を示す回答が若干数得られた。

- 「お、きたきた！とうれしく感じた。」
- 「楽しいので、もっと集めたいなと思った。」
- 「この光を渡してなるものかと思った。」

以上の回答から、クリーチャが集まる様子が、被験者らに何らかの影響を与えていることが分かる。

**質問 3：**光が自分のところから離れていくときに何か感じましたか？

次のように、「寂しい」、「残念」、「悔しい」、「悲しい」、「疎外感」などの感情を記述した回答が多数（10名）得られた。

- 「さみしく感じた。」
- 「自分の話が途切れてしまったかと少し焦った。」
- 「集まってきたものが離れていて若干の寂しさを感じた。」
- 「あー、行ってしまうという残念な気持ち。」
- 「何かくやく感じた。」
- 「なんだか悲しくなった。」

また、次のように発話を継続させたいと思ったことを示す回答が得られている。

- 「必死に映像を思い出そうとした。他人の光を何とか奪おうと思った。」
- 「発言が少ないので、発言しなくてはと感じた。」
- 「もっと話し合いに貢献しなければ！と思った。」

以上の回答から、クリーチャが離れていく様子が、被験者らに何らかの影響を与えていることが分かる。

**質問4：**その他、実験について気になったことや感想

- 「普段体感できないことなので楽しかった」
- 「意外と積極的に話せた自分に驚いた」
- 「活発に意見を出すことに面白さを感じた」
- 「光が集まってくるので、ついついしゃべってしまった」

以上が感想として多く得られた回答である。

本実験の結果から、会話に介在する TableTalkPlus のクリーチャの動きやその変化が会話参加者に認識され、「聞き手性」の表示をアシストするものとして、あるいは自分から聞き手の関心が離れていくような感覚を与えるなど、「会話の場」の認識に何らかの影響を与え得ることを確認した。

## 2.6 実験 (2)

### 2.6.1 目的

本実験は、実験 (1) の結果を受けて、TableTalkPlus のクリーチャが実際の会話にどのような影響を与えるのか検証することを目的とする。

### 2.6.2 被験者

男性 17 名、女性 7 名、計 24 名 (20–24 歳、平均 21.1 歳) が実験に参加した。被験者は互いに顔見知りである 3 人 1 組の会話集団 (全 8 組) に組み込まれた。なお、全ての被験者は、TableTalkPlus を使用しての会話は初めてであった。

### 2.6.3 計画

2.3 で述べた Crowd を使用し、クリーチャが発話に対して反応する速度を条件とした。本実験では、次に示す 3 つの条件を用意し、1 分ごとにそれらの条件を切り替えるという手法をとった。なお、カッコ内に示す数字は一つのマイクだけに入力がある状態で、クリーチャが発話者のもとに集まりきるまでの時間を示している。

**条件 A** 高速 (1.7 秒)

**条件 B** 中速 (2.3 秒)

表 2.2: 質問紙 (4項目, 自由記述)

質問番号	質問内容
1	実験中, 光の動きに何か違いを感じましたか?
2	光が自分のところに集まってきたときに何か感じましたか?
3	光が自分のところから離れていくときに何か感じましたか?
4	その他, 実験について気になったことや感想など

### 条件 C 低速 (2.7 秒)

これらは順序効果を考えて B, A, C の順で遷移するものとし, 2 周 6 分を 1 セッションとした。また, セッション中クリーチャの速度が変化することは, 被験者には教示しなかった。

### 2.6.4 手続き

実験は Crowd の実験用プログラムによる映像をテーブル上に投影しながら, 被験者に自由会話を行ってもらおうという手順で行った。被験者はまず, 円滑な会話を誘引するきっかけとして, ビデオ視聴ブースで 5 分程度のピングーのビデオを見せられる。次に TableTalkPlus が設置された会話ブースに移りテーブルの周りに着席し, TableTalkPlus に接続されているマイク型マイクを装着した後, 実験者の合図によって 6 分間の自由会話を行った。なお直前に 1-2 分程度, インタラクションの様相について確認するための練習時間を設けた。また, 互いの顔よりもテーブル上のコンテンツを見やすくするために, 実験 (1) よりも部屋の照明を暗くした。6 分間の会話が終了した後, 自由記述形式の質問紙 (表 2.2) に回答し, 実



験を終了した。

なお、本実験の質問紙の質問内容（表 2.2）が実験 (1) のもの（表 2.1）と同様ではあるが、それぞれ実験条件が異なるため、別の回答が得られるものと予測される。

### 2.6.5 結果

結果は、クリーチャの速さに対する被験者の振舞いによって評価した。評価項目は平均発言時間（総発言時間を発話回数で割ったもの）、発話回数、6分間で被験者の発話した割合（3人の合計値）とした。なお、全ての発言内容から相槌や「笑い」といったバックチャネルの要素を除いたデータを用いて評価した。

全体の被験者 24 名の内、クリーチャの移動速度によって、評価項目として定めた平均発言時間や発話回数に影響が認められた被験者は 9 名であった。その振舞いの内訳は、クリーチャの移動速度が増すごとに平均発言時間が減った被験者が 4 名、発話回数が減った被験者が 2 名、平均発言時間・発話回数が共に減った被験者が 1 名、逆に平均発言時間が増えた被験者が 1 名、発話回数が増えた被験者が 1 名である。また影響が認められた 9 名のうち、5 名がグループ内で最も発話の割合が多かった被験者であった。

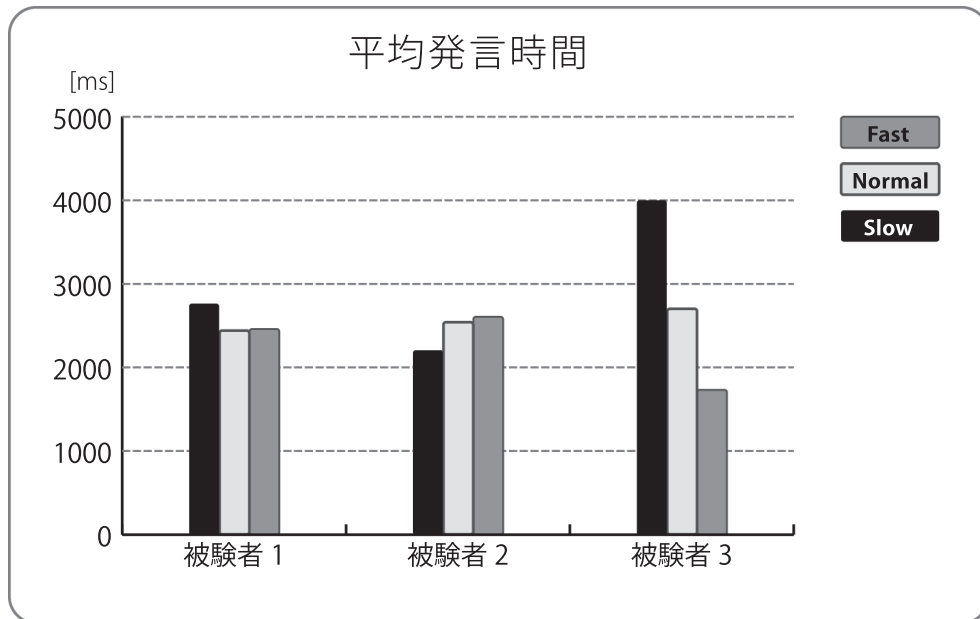


図 2.5: 平均発言時間

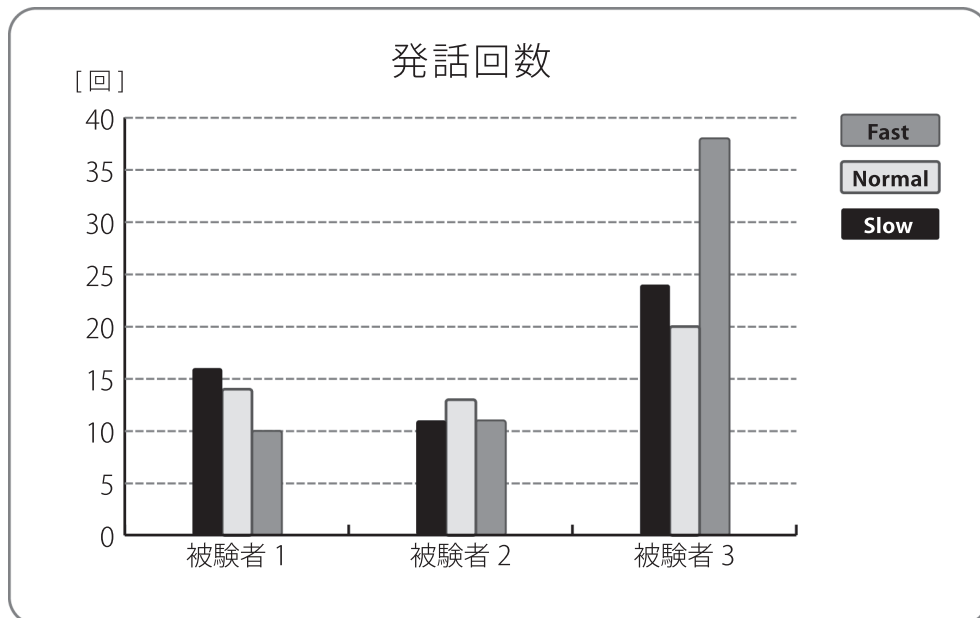


図 2.6: 発話回数

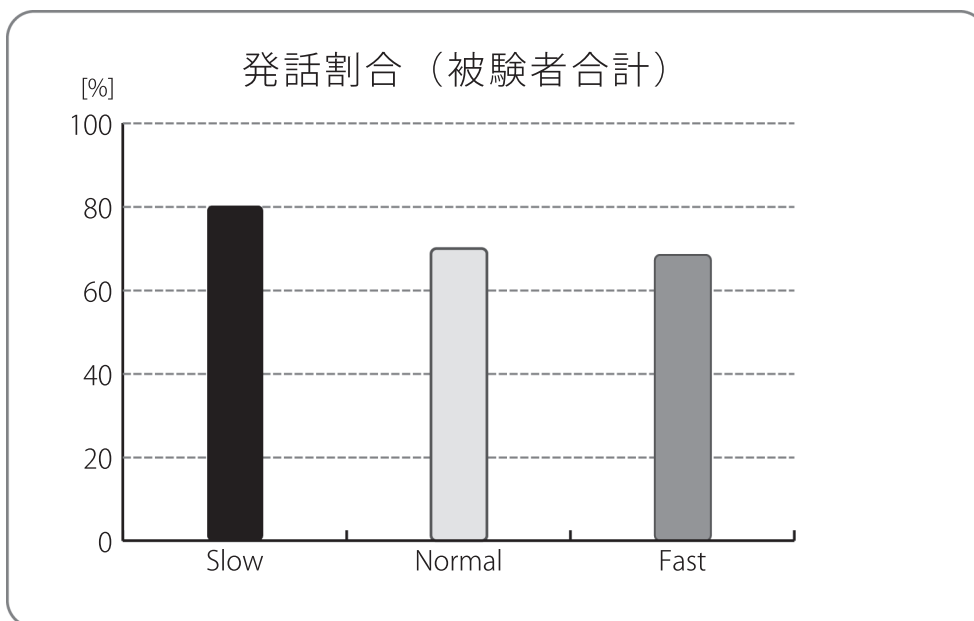


図 2.7: 発話割合 (被験者合計)

図2.5, 図2.6, 図2.7に代表的なグループ<sup>4</sup>における評価値をグラフ化したものを示す。

図2.5, 図2.6のグループの場合, 特に被験者3について, クリーチャの移動速度の変化が, 発話の継続時間や回数に大きな影響を与えていることが分かる。また被験者3は, グループ内で最も発話の割合が高いことが分かる。さらに図2.7から, クリーチャの移動速度が速くなるにつれて, 被験者らの発話割合が徐々に減っていることが分かる。このように, 発話の継続時間や回数, 発話の割合に着目すると, クリーチャの移動速度の変化が特定の被験者の発話行動に影響を与えていることが分かった。

つぎに, 被験者全体の傾向を統計的な手段を用いて調べた。ここでは, クリーチャの移動速度の変化が被験者の発話におけるパワー (音量) の調整に与えた影響を探った。被験者の発話が35dB<sup>5</sup>を超えたフレーム数をカウントし, クリーチャの移動速度毎に整理したものを表2.3に示す。ここで1フレームは1/40秒 (0.025秒) である。表2.3から, 被験者によってフレーム数に差があることが分かる。これは声量の大きい被験者がいる一方で, 静かに発話を行う被験者もいるということを示している。さらに興味深い点は, クリーチャの移動速度の変化に従って, 被験者の発話 (フレーム数) も変化している点である。この結果 (表2.3) について, ANOVAにより分散分析を行った。表2.4にその結果を示す。

表2.4から, クリーチャの移動速度の変化と被験者の35dBを超える発話のフレーム数との間に有意差が確認できる ( $F=5.666$ ,  $df = 2/46$ ,  $p<0.01$ )。さらに, 表2.4の結果に対して下位検定を行った。表2.5にその結果を示す。表2.5に示されるとおり, Ryanの方法による多重比較を行った結果, クリーチャの移動速度の「低速」と「高速」の間に5%水準で有意な差が認められた。以上を考慮すると, クリーチャの移動速度の急激な変化が被験者の発話におけるパワー (音量) 調整に影響を与えていることが分かる。

<sup>4</sup>全体のデータをとるとグループレベル・個人レベルでみられる特徴量が平均化されてしまうため1つのグループに絞る。

<sup>5</sup>35dBは被験者全員の平均の発話音量 (パワー) をとった値である。

表 2.3: クリーチャの移動速度に対する 35dB を超える被験者の発話のフレーム数

	低速	中速	高速
被験者 1	2220	2038	1926
被験者 2	2357	2749	2278
被験者 3	2641	2844	2364
被験者 4	3099	2923	2578
被験者 5	3218	2988	2997
被験者 6	3567	3428	3322
被験者 7	4055	3837	3337
被験者 8	3928	4157	4001
被験者 9	4345	4367	4177
被験者 10	2498	2201	2661
被験者 11	2488	2265	2445
被験者 12	2597	2396	2685
被験者 13	1360	638	925
被験者 14	2002	1587	1834
被験者 15	2161	1805	1713
被験者 16	1975	1831	1251
被験者 17	2111	1962	1828
被験者 18	3017	2437	2562
被験者 19	3390	3678	3736
被験者 20	2880	3356	3051
被験者 21	3307	3551	3438
被験者 22	2716	2601	2740
被験者 23	2885	2423	2349
被験者 24	3237	3134	3142

表 2.4: 分散分析表

	偏差平方和	自由度	平均平方	分散比 (F 値)	p 値	有意差判定
移動速度	469926.3	2	234963.2	5.666	0.0063	**
誤差	1907485.0	46	41467.1			
合計	45794175.5	71				

表 2.5: Ryan の方法による多重比較の結果

比較する群	群数	多義的有意水準	t 値	p 値	有意差判定
低速-高速	3	0.0166667	3.341	0.0016625	有意差あり
低速-中速	2	0.0333333	2.026	0.0486118	有意差なし
中速-高速	2	0.0333333	1.316	0.1948463	有意差なし

最後に、質問紙の各質問項目ごとに多く見られた回答や特徴的な回答を示す。簡単な考察とあわせて以下に述べる。

**質問 1：**実験中、光の動きに違いを感じましたか？

この質問に対して、「違いはわからなかった」と回答する被験者は 18 名（24 名中）であった。本実験では、明らかにクリーチャの移動速度を変化させているが、被験者らはその違いを見分けられなかったことがこの回答から明らかとなった。

また、次のようにクリーチャの動きの法則について示す回答が得られた。

- 「声の大きさや長さによって集まる光の数が違ってすごいと思った」
- 「声量や話し方で集まり具合が違う？」
- 「何の法則で集まるのか不思議に思った」

以上の回答から、シンプルなルールをもったクリーチャが、筆者らの狙い通り、集団として複雑に動いていたことが明らかとなった。

**質問 2：**光が集まってきたとき何か感じましたか？

次のように自分が話し手となっていることを自覚し、会話を独占しようとする回答が得られた。

- 「嬉しかった。注目されていると感じた」
- 「自分のところにとどめておきたいと思った」

以上の回答は、クリーチャの集まり方が、発話の継続時間に影響を与えている可能性を示唆するものである。

また、次のようにクリーチャの数が 3 人で均等になることを要求する回答が得られた。

- 「集まりすぎると皆に分けたくなる」
- 「3 人に同じくらいあるようにしたいと思った」

以上の回答は、TableTalkPlusのコンテンツを人為的に操作し、被験者らの発話の割合を調整できる可能性を示唆するものである。

**質問3:** 光が離れていったとき何か感じましたか？

次のように体の動きに変化が見られる回答が得られた。

- 「普段よりも“聞く”姿勢が強まった」
- 「光を見ると自然と話者の方へ向いていた」

以上の回答は、クリーチャの離れ方が、発話者の行為に何らかの影響を与えている可能性を示唆するものである。

**質問4:** その他、気になったことや感想

- 「楽しかった。光ばかり見てしまった」
- 「暗くて会話が弾まないかと思ったが、慣れると問題なく普段のように会話することができた」

以上が質問紙への回答として得られた結果である。

## 2.7 考察

私たちの会話は、相手に何かを伝えたり、何かを依頼するなどの情報伝達・行為遂行的な要素だけではなく、その場を一緒に共有したり、その場の秩序を維持することを介して、お互いの社会的なつながりを維持したり、調整しあうという側面がある。また、会話の中の聞き手、話し手、傍観者などの役割は、会話参与者の中で相互行為的に組織される。それぞれの参与者は、この参与フレームにおける相互行為秩序に取り込まれながら、相手からの語り掛けに応答責任を担ったり、「あなたの話をちゃんと聞いていますよ」といった聞き手性 (hearership) の表示が欠かせない。こうした、会話参与者の拮抗した関係性やダイナミクスは、お互いの行為を制約するような社会的な「力」を備えるもので、「会話の場」と呼んできた。



本章において議論した TableTalkPlus は、これまで会話の中で「地」としてあった、参与者同士の拮抗した関係性やダイナミクス、それを支える聞き手性などの情報を「図」として 顕在化させることを狙いとした。

実験(1)では、会話参与者(被験者)が TableTalkPlus の存在やそのクリーチャの反応をどのようなものと認識したのかを被験者の質問紙への回答を手掛かりに調べた。その結果、クリーチャたちが自分のところに集まる状況に対しては、「自分が話し手であることを再認識する」ことを示す感想が得られている。また、クリーチャが離れていく状況に対しては、「他の参与者の注意が自分から離れていくような寂しさ」を感じ、また「自分のところに集めてみたいという気持ち」を引き起こさせている。また、クリーチャを自分のところにだけ集めている状況が続くと、「他の人にも譲らなければ」という気持ちを引き起こさせていることもわかった。これらから、会話の背後にある関係性やダイナミクスを顕在化させるという TableTalkPlus の初期の目標がクリアできていることを確認した。

実験(2)では、会話に介在する TableTalkPlus 中のクリーチャの移動速度を人為的に変化させることで、その参与者の発話時間や発話回数、あるいは会話全体における発話の割合など発話制御に与える影響を調べた。その結果、クリーチャの移動速度の変化に影響を受けやすい被験者と影響を受けにくい被験者があった。その影響を受けやすい被験者の例においては、クリーチャの速度に合わせて、発話継続時間が短くなり、結果として発話回数などが増えている。クリーチャの動きに合わせて、発話を制御していることを確認できた。一方、人為的な操作による秩序の乱れを参与者全体で修復する動きもあり、全体の平均を取るとその中に埋没する傾向もみられた。これはクリーチャの動きの影響を相殺するように相互行為秩序が機能しているためと考えられるが、今後、詳細に調べていく必要がある。

以上の結果から、多人数会話における参与者を媒介する TableTalkPlus の役割や意義が幾つか明らかになった。

その一つは、会話の中で「地」としてあった参与者の拮抗した関係性やダイナミクスを「図」として顕在化する働きがあり、会話の性質を情報伝達の・行為遂行的な側面に加え、会話の場を維持したり、調整することに意味のある自己目的的

な側面を強める効果が期待できる点である。被験者からの一部のプロトコルや感想に見られたように、会話を楽しくさせたり、盛り上がらせるような役割や、人と人とのつながりを媒介するソーシャルメディアとしての応用が期待できる。

また、実験(2)で示されたように、人為的にクリーチャの移動速度などのパラメータを統制することで、個々の発話者の発話制御を含め、会話全体にも影響を与え得ることを確認した。これはTableTalkPlusのような媒介物を会話の中に介入させることで、会話の場そのものの第三者によるデザイン可能性を示している。

もう一つは、一種のプローブ (prover) として、会話における相互行為秩序の研究への応用可能性である。会話分析研究やエスノメソドロジーでは、研究者は傍観者の立場からその会話を観察し、分析・記述するという手法を取ってきた。また、文化人類学や社会学では、その会話に一人の参加者として参加し、参加者の立場から分析・記述する参加観察の方法などを用いている。本章で提案し、実装したTableTalkPlusは、もう一つの研究手法を提供できる。つまり、その会話に媒介物として介入し、そのパラメータを統制しながら、その場の性質を明らかにするという構成的な方法論である。実験(2)はその可能性を一部確認したもので、今後は「会話の場」における相互行為秩序の生成基盤等について、更に詳細な研究を進めていく必要がある。

## 2.8 おわりに

本章では、会話の背後にある参加者間の関係性やダイナミクスに着目し、CGを用いて、その視覚化を試みるTableTalkPlusのコンセプトとその実装について述べた。

実験結果から、話し手と聞き手との関係の中で揺れ動く、TableTalkPlusのクリーチャの動きが会話参加者の発話行動を制約し方向づけるような、アーティファクトとして機能することを示してきた。また、考察においてTableTalkPlusのソーシャルメディアとしての応用可能性や会話研究への応用可能性について議論した。

今後は、この研究で使用した Crowd だけではなく、さらに効果的なコンテンツのデザインを探る予定である。また、コミュニケーション障害児に対するコミュニケーション支援環境やコミュニケーションスキルの学習支援環境など、具体的な応用について検討を進めたいと考えている。



## 第3章 説得性の社会的構成に着目したメディア技術—タイプII

### 3.1 まえがき

人の身体を揺り動かしたり、その行為を促すようなインタラクションにはどのようなものがあるだろう。一つには、あるモノを押すことで、それを物理的に移動させたり、倒したりするように、相手の身体を直接に押ししたり、引いたりする「物理的なインタラクション」が考えられる。また、「おばあちゃん、お薬の時間だよ!」というような言葉を掛けることで、その相手にある特定の行為を促すこともある。これは物理的なインタラクションに対して、「情報的なインタラクション」と呼べるものである。

これらのインタラクションのタイプに加え、日々の生活の中では、ある状況や場に置かれたとき、思わず行為が促されてしまうということもある。例えば、なにか困っている人を見かけたときに、思わず助けようと身体が動いてしまう。あるいは、知り合いから挨拶されたときになど、思わず応答責任のようなものを感じてしまう。集団での生活の中にあっては、他の人たちの頑張る様子に、自分の行動も何かしらの影響を受けてしまうということもあるだろう。そこでは、物理的な力を介したインタラクションでも、他者への指示を伴う言葉の力を介したインタラクションでもない、人を思わず揺り動かしてしまうような、ある特定の「場」によって引き起こされるインタラクションが存在するように思われる。

これらの感覚は、日々の生活の中で経験的に感じることもあるものの、人の身体を思わず揺り動かしてしまう「場」とはどのようなものなのか、それはどのようにして生じるものなのかについての議論はまだ十分なものではない。

そこで筆者らは、この人の行為を思わず促すような「場」を人為的に生みだしながら、このような「場」を支えている背後にある原理を探ることはできないかという観点から、いくつかの構成的な研究を進めてきた [79, 103]。また、このような人を思わず揺り動かす「場」の創成技術やハンドリング技術の応用可能性を探っている。

これまでの研究では、これらの試みの一環として、スウィングング・ルーム (swinging room) 中での人の揺動運動を調べた Lee らの Optic Flow Field に関する研究を手掛かりに、この Optic Flow Field のアイデアを社会的相互行為の場に展開し、その効果を確認した。

具体的には、TableTalkPlus という研究プラットフォームを構築し、Lee らの試みにある光学的流動のアナロジーとして、会話における会話参与者間の関わりの結果として生まれるクリーチャの群れを CG によって生成した。また、これらのプラットフォームを使用し、会話の中で生まれるクリーチャの群れに対する人為的な操作が会話参与者の会話調整にどのような影響を与えるのかを調べた。

本研究では、この TableTalkPlus の拡張として NEXUS+ というプラットフォームを構築し、会話における会話参与者間の関わりの結果として生まれるクリーチャの群れの動きを、ネットワークで接続された2つの離れた会話の場に展開することを試みた。

ここでは、2つのパーティ (= 会話の場) の間での会話の盛り上がり反応して、2つのテーブル間を行き来するクリーチャの群れの動きを生成し、その動きがお互いの会話行動にどのような影響を与えるのかを調べた。その結果、Lee らの Optic Flow Field に関する実験と同様に、ネットワークを介して行き来するクリーチャの群れの動きが2つのパーティにおける会話参与者の発話行動に影響を与えるものであることを確認した。

本論文では、これらの内容を次の構成に沿って詳しく述べることにする。本研究の背景、狙い (3.1 節) に引き続き、生態心理学における Optic Flow Field に関する研究とその社会的相互行為に展開することの意味を整理し、本研究のプラットフォームの一つである TableTalkPlus の特徴を説明する (3.2 節)。その後、も

う一つのプラットフォームである NEXUS+ の構成 (3.3 節) とそれに基づく実験と考察 (3.4 節) を述べ、最後に総合考察とまとめを述べる (3.5 節)。

## 3.2 Optic Flow Field とその社会的相互行為の場への展開

### 3.2.1 視覚的運動制御

人の身体を揺り動かすような「場」の存在について、生態心理学の分野では、Optic Flow Field の研究 [47, 48] が知られている。

この研究では、光学的流動 (optic flow) を人為的に動かすことで、その動きに合わせた人の身体の揺動運動を誘発できることを明らかにした [108]。具体的には、被験者を取り囲む部屋の壁を前後に揺らすことの可能な「スウィングルーム (swinging room)」という実験環境を構築して、被験者の視点に近づく、あるいは遠ざかるような光学的流動を人為的に作り出した。その実験の中で、被験者の視点に近づくような光学的流動は、その被験者の姿勢を前のめりにさせる。また、その視点から遠ざかるような光学的流動は、その姿勢を反らせるような動きを引き出すことを確かめた。

この研究では、私たちの身体の姿勢は、脳からの指示により制御されているのみならず、むしろ身体の動きとそれを取り囲む環境との関わりから生みだされた情報によって支えられていることを示した。身体の動きに呼応する見えの変化として、いわゆる光学的流動を生みだし、それを手掛かりに姿勢を制御しているというわけである。これは J.J.Gibson により「視覚的運動制御」と呼ばれている [24, 100]。

また、この Optic Flow Field のような、思わず身体を揺り動かしてしまう「場」というのは、身体とそれを取り囲む環境との間にある行為と知覚との連続的なカップリングによって生みだされたものと考えられる。

### 3.2.2 他者（社会的な環境）との関わりへの展開

私たちの身体と環境との関わりに限らず、他者という社会的な環境との関わりの中でも、自分の行為が思わず引き出されてしまうという場面を経験する。その典型は、会話などの社会的相互行為の場面においてだろう。

知り合いから挨拶されたとき、その挨拶を無視して立ち去ることは難しい。ここでは、思わず応答責任のようなものを感じてしまう。あるいは、会話中での沈黙では、「ここで何かを話さなくては…」という気持ちになる。会話の中でも、それぞれの会話参加者の行為を促すような「場」が存在し、そこで発話交替のダイナミクスを生みだしたり、会話連鎖を組織化していく動因となっているように思われる [114]。

生態心理学の分野においても、身体と環境間における行為-知覚のカップリングに関する議論をこの社会的な環境に展開する試みが行われている。その一つは、個人間協調 (inter-personal coordination) に関する一連の研究である。例えば、Kugler と Turvey は、二人の被験者に、片手に一つずつ振り子を持たせ、手首を軸として振らせるような実験を行った。また、ここで生じた肢体間協調の現象を説明するダイナミカルなモデルを提案している [43]。

ただ、これらの議論は身体レベルでのリズム運動に限られており、そのまま日常会話などの社会的相互行為の領域での議論に展開することは難しい。会話場面でのやり取りされる情報は、主にシンボル化された言語であり、離散的なものである。そのため、その背後にある「場」の存在は、言語的なシンボルのやり取りの中に隠れやすいものであった。

### 3.2.3 Optic Flow Field の社会的相互行為への展開

本研究では、Lee らの Optic Flow Field の実験における光学的流動を手掛かりに、会話の場におけるダイナミクスを近似的に表現することを試みた。

視覚的運動制御における光学的流動の働きを考えると、そこで重要な役割を果たしているのは、その連続的な流れである。自分の身体の動きに協応した見えの



変化(=光学的流動)によって、自分自身の動きやそのスピードを自覚できる。これは「生態学的な自己 (ecological self[97])」と呼ばれるものである。同時に、引き続きそのような状態に移行していくのか、あるいは壁にぶつかるまでの時間 (time to contact[68]) などの予期的情報を提供してくれる。

日常会話においても、会話参与者間での発話交替のタイミングにおいて、「ターン移行適切場 (turn transition relevance place)」の存在が指摘されている [73][29]。「そろそろ相手のターンをこちらに移行してもいい頃だろうか」というタイミングを図る手掛かりとなるもので、これは「壁にぶつかるまでの時間 (time to contact)」と同様に、話し手の発話やイントネーションの変化などのさまざまな手掛かりを利用して予期可能なものである。

また、何か話そうとするとき、その相手がターンを譲ってくれない、あるいは容易にターンを譲ってくれたという、会話参与者間での関わりの中で、その中で自分の立場や存在の大きさを認識できる。これは「生態学的な自己 (ecological self)」に対して、「対人的な自己 (social self[97])」の獲得と呼ばれるものである。

このような会話の中での関係性や状態の変化を連続的に表現する手段として、CGを利用したクリーチャの群れの動きとして生成し、その会話参与者の取り囲んでいるテーブル上に投影する TableTalkPlus[106] というインスタレーションを構築してきた。

### 3.2.4 TableTalkPlus とその特徴

この TableTalkPlus の生成するクリーチャの群れの動き (図 3.1) は、次のような性質を持つ。(1) 誰も発話していない状態では、それぞれのクリーチャは離れた状態で自由な動きをしている。(2) 誰かが話しはじめると、クリーチャの群れはその話し手の方に移動し、そこに集まる。また、話し終わると、そのクリーチャの群れは、それぞれ離れた状態で自由な動きをはじめ。(3) 誰かが話している途中で、他の誰かが話しはじめると、その群れの一部は新たな話し手の方へ移動し始める。

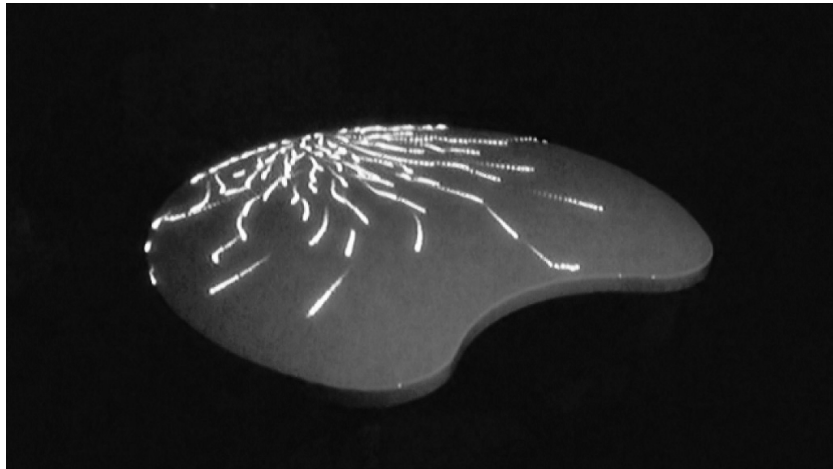


図3.1: Table Talk Plus の動作例

これを Optic Flow Field の議論と対応づけると、話し手は自分の方に移動するクリーチャの群れを見ながら、自分が発話を継続していること、そして他の参与者に話し手として承認されていることを自覚できる。

例えば、(A) より多くのクリーチャが自分の方だけに集まってくる様子を見て、自分の話し過ぎを心配して発話を終了したり (図3.2)、(B) より多くのクリーチャが他者の方だけに集まっていく様子を見て、相手の会話を横取りすることを心配して発話を躊躇したり (図3.3)、(C) 発話のタイミングを皆で調整してテーブルの中央にクリーチャを留めておこうとする振る舞い (図3.4) は、クリーチャの群れの動きが発話交替のタイミングを図る手掛かりとして、また、その会話の中で自分の立場や存在の大きさを自覚する手掛かりとして使用された例といえる。



図3.2: より多くのクリーチャが自分の方に集まってくる様子を見て発話を終了する

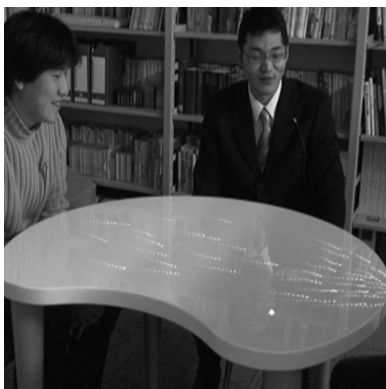


図3.3: より多くのクリーチャが他者の方に集まっていく様子を見て発話を躊躇する



図3.4: 発話タイミングの調整によりテーブル中央にクリーチャを留めようとする

### 3.3 NEXUS+

本研究では次に、この TableTalkPlus のクリーチャの群れの動きを、ネットワークで接続された2つの離れたパーティ (=会話の場) に展開するために、NEXUS+ というプラットフォームを構築した。

#### 3.3.1 システム構成

NEXUS+は、2つの離れたパーティ間での会話の盛り上がりに対応して2つのテーブル間をクリーチャが行き来するインタラクティブメディアである。ここでは、3人で行われる多人数会話を一つのパーティとし、2つのパーティで1組のNEXUS+を使用することとした。

#### 3.3.2 ハードウェア

一つのパーティで使用するハードウェアの構成は Synchronized Heart Beat System[92] や Crossing Colorful Communications[63] などのテーブルウェアの枠組みに位置付けられる。ここでは、会話参加者の音声を取得するための3つのワイヤレスマイクと、3つのマイクレシーバ、これらをPCに接続するためのUSBオーディオインタフェース、CGの生成及びデータ通信用のPC（クライアント機またはサーバー機）、CGを投影するための液晶プロジェクタと映像の投影対象となるテーブルで構成している。

この構成を1セットとし、クライアントPCのセット（図3.5左半分）とサーバーPCのセット（図3.5右半分）をそれぞれ別々の部屋に設置して、互いをネットワークケーブルで接続することで、NEXUS+は一つのインタラクティブシステムとして機能する（図3.5）。

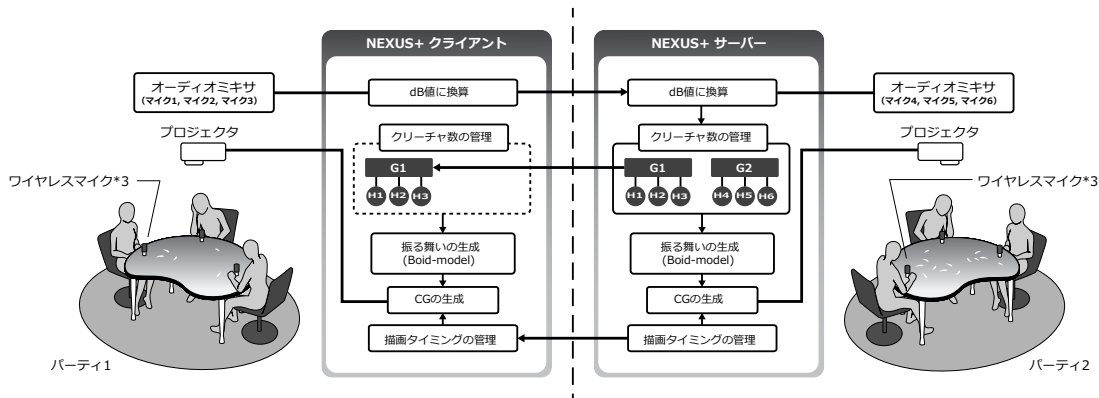


図 3.5: NEXUS+のシステム構成と内部の処理

### 3.3.3 ソフトウェア

クライアント PC およびサーバー PC 内の処理は (1)MAX/MSP[88] を用いて、クライアント PC とサーバー PC は会話参与者一人一人の保有するワイヤレスマイクから得られる音声のパワーを dB 変換し、3 人分を合計する。(2) この合計値をクライアント PC は UDP/IP ネットワーク通信によってサーバー PC へ送信し、サーバー PC は収集した各パーティの合計値をもとに、パーティ間の比を計算する。この比に従って (3) サーバー PC はそれぞれのパーティに表示するクリーチャの数を計算し、その値をクライアント PC に伝える。(4) 両パーティの描画タイミングが一致するように、クライアント PC とサーバー PC は同期処理を行いながら、Processing[96] を用いてクリーチャ数に応じた CG アニメーションをそれぞれ生成する。(6) 生成されたアニメーションは PC からプロジェクタへ出力され、それぞれのテーブル上に投影される。

#### 3.3.4 インタラクションデザイン

このテーブル上に表示されるクリーチャの量の変化を連続的なものとして表現するために、次の 2 点に注意してアニメーションを生成している。

一つは、会話における発話の量的な変化に着目して、パーティ間のクリーチャの

相対的な量を変化させている。まず、1秒間に30回の周期で各ワイヤレスマイクから得られるパワーの合計値を更新し、その時々合計値を2秒間分(60個)足しあわせる。この2秒間分の合計値を2つのパーティ間で比較し、その比に従って、各パーティのテーブル上に表示するクリーチャの数を決定している。

例えば、あるパーティAが一人ひとり30dB以下で継続的かつ同時に発話している状況で、別のパーティBが一人一人70dB以上で継続的かつ同時に発話している状況を作ると、10秒後にはパーティAとパーティBのクリーチャは20匹:40匹に分散し、20秒後には15匹:45匹に分かれる。このようにすることで、急激にクリーチャの量が増えることが避けられ、また、2つのパーティの会話の状態に応じて相対的に変化させることができる。

もう一つは、クリーチャの量を増減させる場合の処理である。あるパーティのクリーチャを増加させる場合はテーブルの中央にある仮想的な穴から、こちら側にクリーチャが移動してくるというアニメーションを表示させている。また、あるパーティのクリーチャを減少させる場合はテーブル上を自由に移動していたクリーチャを中央の仮想的な穴から消していくというアニメーションを表示させている。このようなアニメーションを加えることで、徐々にクリーチャが移動していく様子を見て取れるようにした。

こうした流れが連続的に起こることで、会話の静かになった部屋のテーブルでは、徐々にクリーチャが減っていき、自分たちが発話を控えているということ、そして、他のパーティよりも落ち着いたある会話をしていることを自覚できる(図3.6)。それとは対照的に、会話の賑やかな部屋のテーブルには、次第にクリーチャが集まっていき、自分たちが発話を継続していること、そして、他のパーティよりも活発に発話していることを自覚できる(図3.7)。

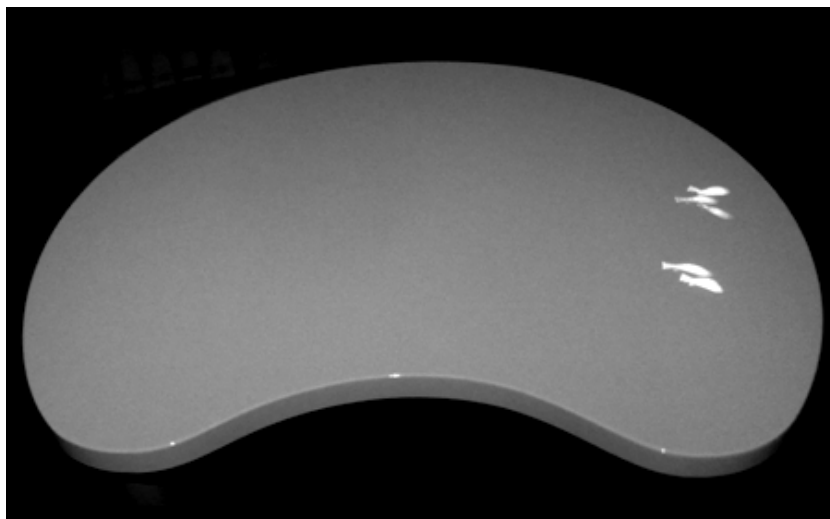


図 3.6: NEXUS+の動作例 1

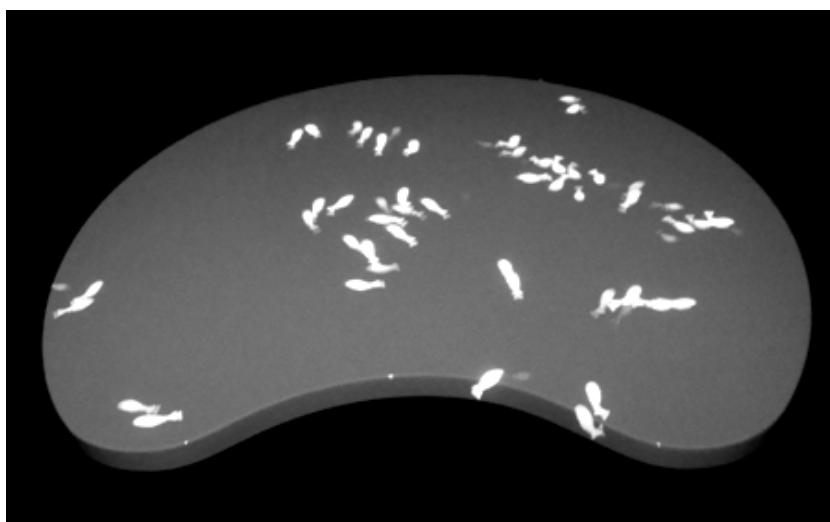


図 3.7: NEXUS+の動作例 2

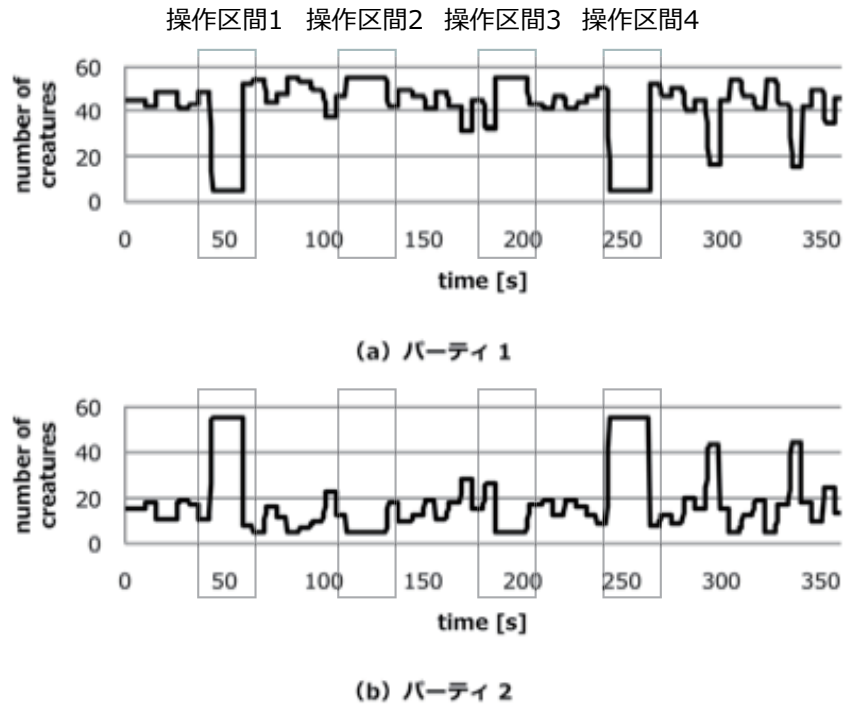


図3.8: 各パーティのクリーチャ数の時間的な変化と操作区間

ここでは、この2つのパーティ間を行き来するクリーチャの量を、会話の盛り上がりとは別に人為的に変化させる区間を設けた。図3.8は各パーティのクリーチャの量の時間的な変化を示したグラフである。ここでは、会話の開始から約50秒後、120秒後、190秒後、260秒後の4つを操作区間の開始点とし、約25秒間、クリーチャの相対量を5匹:55匹（もしくは55匹:5匹）に人為的に変化させている。図3.8に示すとおり、これら4つの操作区間において、クリーチャの相対量が急激に変化していることが分かる。

このようにネットワークを介して行き来するクリーチャの量を操作することで、Optic Flow Fieldに関する議論をネットワークを介した社会的相互行為の領域に拡張する実験を行った。



## 3.4 実験

### 3.4.1 目的

本実験では、2つの会話の場の関係性に合わせて変化するクリーチャが、会話参加者の発話制御にどのような影響を与えるのかを、各テーブルに投影されるクリーチャの量を人為的に操作し変化させることで調べた。

具体的には、会話の盛り上がりに対応するクリーチャが別々のパーティ間を行き来することで(1)相手との一体感、つながった感じを引き出すことはあるのか、(2)別々のパーティがネットワークを介して一つのクリーチャを共有することで向こう側のグループの状況や性格が自分たちのグループにも伝わることはあるのか、(3)そのクリーチャの人為的な変化の影響はどのような被験者のどのような発話行動に現れるのか、の3項目を検討することを目的とした。

### 3.4.2 被験者

男性18名(20–25歳, 平均21.9歳)の大学生が実験に参加した。被験者は相互に顔見知りである3人一組の会話集団(全6パーティ)に組み込まれた。その内、別々の部屋で同時に会話を行う2パーティを選び、これを1組として全3組に振り分けた。なお、全ての被験者はNEXUS+を使用しての会話は初めてであった。

### 3.4.3 手続き

実験の流れを以下に示す。(i)ビデオ視聴ブースにおいて「ビデオの内容を覚えてストーリーを正確に話す」という課題(共同想起対話[16])を被験者に伝えてから、5分程度のビデオ<sup>1</sup>を見せる。(ii)NEXUS+が設置されたそれぞれの部屋に移動し、体験セッションを5分間行う。(iii)本セッションを6分間行う。(iv)半構造化

---

<sup>1</sup>Hanna, W., Barbera, J., Jones, C.: *TOM & JERRY: Sci-Fi & Fantasy type episodes*[DVD], Warner Home Video (2003).

インタビューを行う。

なお、体験セッションでは別々の部屋にクリーチャが行き来することを確認してもらった。その手順として、システムが設置された部屋をビデオチャットにより接続し、互いに別々の部屋のテーブル上のクリーチャが見える状況を作った。また、この時にクリーチャは「盛り上がっているテーブルの方へ移動します」とだけ伝えた。そして、そのように動作することを実際に会話を行いながら被験者に体験してもらった。

一方、本セッションでは、ビデオチャットの接続を切り、互いのテーブルの状況が分からない状態で行っている。また、会話中はこのアニメーションを見ながら会話をするようにと教示した。ただし、このアニメーションをどのように使用するかについては一切教示しなかった。なお、操作区間では、一方のクリーチャの数を5匹、もう一方を55匹に人為的に変化させている。

#### 3.4.4 結果

実験結果の分析には、被験者の発話行動を記録した客観的なデータと半構造化インタビューへの回答という主観的なデータの2つを用いた。

##### 主観評価

まず、半構造化インタビューの結果について、(A)別々のパーティがネットワークを介して一つのクリーチャを共有することで向こう側のグループの状況や性格が自分たちのグループにも伝わることはあるのか、(B)相手との一体感やつながった感じを引き出すことはあるのか、という観点から分析した。

ここでは、半構造化インタビューのなかで多く得られた回答や特徴的な回答について述べる。ただし、カッコ内の記述は筆者による補足とする。

**質問1:**相手のグループが気になることはありましたか?それはどんな時ですか?

次のように、自分たちの会話に変化した時、あるいは、クリーチャの量に変化した時に、相手のグループが気になるという回答が多く得られた(18名中12名)。

- こちらの会話が尽きたとき、魚（クリーチャ）が少ないとき
- 盛り上がっているとき、魚（クリーチャ）が多くいるとき
- 自分のグループの会話が途切れたときに、相手のグループが気になった。魚（クリーチャ）が少ない時

**質問2：**何か話さなければならぬと感じたことはありましたか？それはどんな時ですか？

次のように、会話に沈黙が生じた時やクリーチャの量が減少した時という回答が多く得られた（18名中17名）。

- 一瞬会話が途切れた時、魚（クリーチャ）の量が減った時
- カラフルな魚（クリーチャ）が消えて少なくなった時
- 沈黙になった時、魚（クリーチャ）が向こう（のグループ）に持っていかれた時
- 静かな時に魚（クリーチャ）が消えると、そう思った

**質問3：**話しやすいと感じたのはどんな時でしたか？

クリーチャの量が多い時という回答が8名から、会話の話題が変化した時という回答が5名から得られた。

- 盛り上がっているとき、魚（クリーチャ）が多くいる時
- 魚（クリーチャ）が多く自分の側にいた時

### 人為的な操作の影響が顕著に現れる被験者の特徴を明らかにする

次に、クリーチャ数的人為的な変化に影響を受けやすいグループ（被験者）の特徴を調べた。その方法として(1)操作区間におけるグループ間での発話の平均音

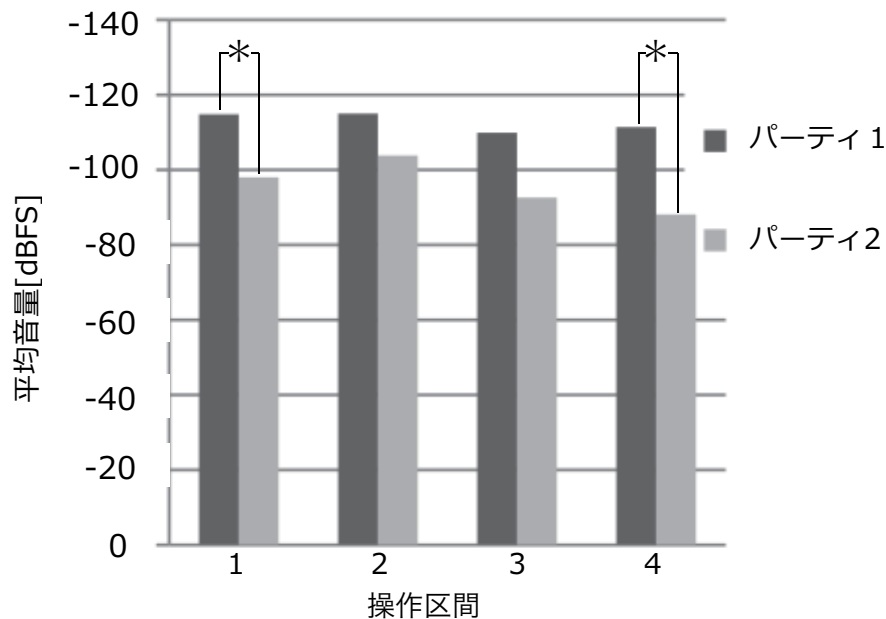


図 3.9: 各操作区間における発話の平均音量 (tukey 検定)

量の違いを分析する (tukey 検定), (2) 操作区間における被験者の発話の質的な変化を分析する, (3) 被験者間の発話の関連性を調べる (コレスポンデンス分析) という 3 つの分析を段階的に行った。

### 発話の平均音量の変化を分析する

すべての組について, 操作区間における発話の平均音量 (3 人の合計値) のパーティ間での違いを調べたところ, 第 1 組目の 2 つの操作区間において優位な差が現れた。その結果を図 3.9 に示す。

この組では 4 つの操作区間の内, 1 つ目 (Tukey(N-1=1, n-1=2)=-3.65730, \*P<0.05) と 4 つ目 (Tukey(N-1=1, n-1=2)=-2.85829, \*P<0.05) の操作区間についてグループ間の平均音量に優位な差が現れた。その理由として, 操作区間とその前の通常区間とでクリーチャの量に急激な差があり, 被験者の発話に影響を与えたことが考えられる。一方で, 2 つ目と 3 つ目の操作区間では, 操作する前の通常区間と操

作区間とでクリーチャの量の変化が少なく、被験者がその違いを見分けることができなかつたため、優位差が現れなかつたと考えられる。

### 発話の質的な変化を分析する

次に、前述の分析において優位な差が現れた組について、操作区間における被験者の発話の質的な変化を分析した。

図 3.10, 図 3.11 は、操作区間における各被験者の頷きや相槌といったバックチャネルをビデオ分析によりカウントした結果である。図 3.10, 図 3.11 より、パーティ 1 の被験者 3 を除いて、クリーチャの増大時より減少時にバックチャネルが増加するという傾向がみられた。ここから、バックチャネルを増加させるというような、相手の発話中に自分の発話を重ねるような発話の制御とクリーチャの減少が関係していることが分かった。

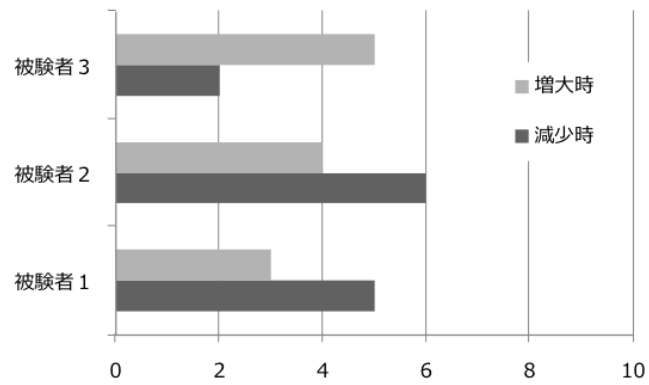


図3.10: 操作区間におけるバックチャネルのカウント数 (パーティ1)

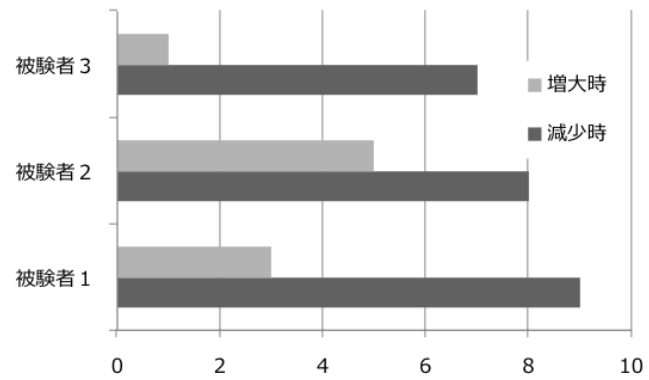


図3.11: 操作区間におけるバックチャネルのカウント数 (パーティ2)

表 3.1: コレスポネンス分析で使ったラベル

ラベル	内容
S1	クリーチャの数が0以上10未満の時
S2	クリーチャの数が20以上40未満の時
S3	クリーチャの数が40以上の時
H1	被験者1のみが発話している時
H2	被験者2のみが発話している時
H3	被験者3のみが発話している時
H1H2	被験者1と被験者2が同時に発話している時
H2H3	被験者2と被験者3が同時に発話している時
H1H3	被験者1と被験者3が同時に発話している時

### 被験者間の発話の関連性を調べる

次に、この組でのクリーチャの量の変化と被験者間の発話の関連性を調べるために、コレスポネンス分析を行った。コレスポネンス分析とは、クロス統計表の中の行と列の各項目の関連性を調べる統計手法である。ここでは、行をクリーチャの量の違い、列を被験者の発話行動の変化としてクロス統計表を作った。

クロス統計表を作成するにあたり、クリーチャの量と被験者一人ひとりの具体的な発話行動を表3.1に挙げたラベルに従ってカウントした。なお、表3.1に該当しない被験者の振る舞いはカウントから除いている。こうして得られたカウント値をもとにコレスポネンス分析を行った。

このコレスポネンス分析結果を散布図で表したものを図3.12、図3.13に示す。このグラフは変数間の類似性を可視化したものである。関連性の強い項目同士はより近い位置にプロットされる。

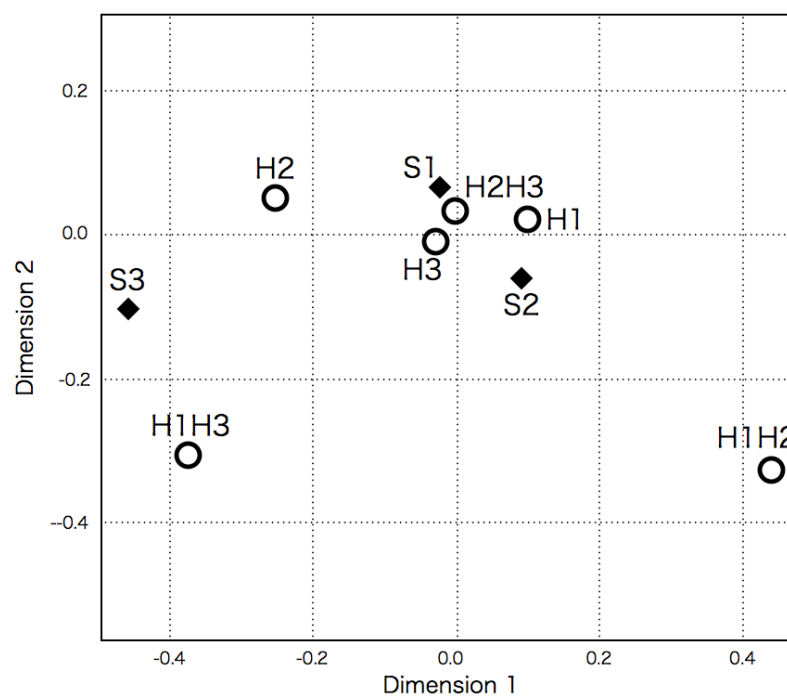


図 3.12: コレスポネンス分析結果 (パーティ1)

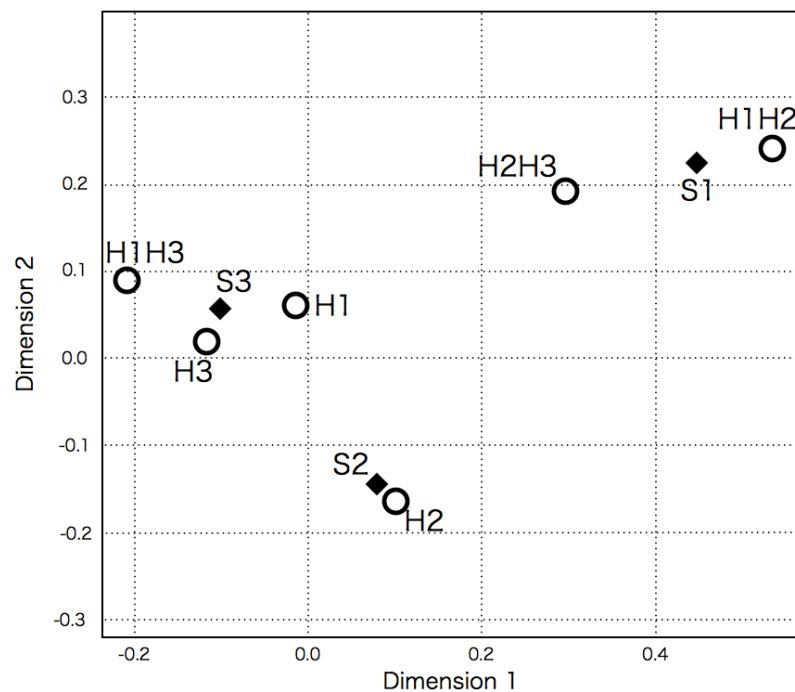


図 3.13: コレスポネンス分析結果 (パーティ2)



パーティ1では、クリーチャの量が比較的少ないS1とS2の時と、被験者1の発話(H1)と被験者3の発話(H3)、被験者2と被験者3の同時発話(H2H3)の関連性が強い(図3.12)。一方、パーティ2では、クリーチャの量が多いとき(S3)と被験者1の発話(H1)と被験者3の発話(H3)、被験者1と被験者3の同時発話(H1H3)の関連性が強い(図3.13)。また、クリーチャの量がS3よりも少ないS2では、被験者2の発話(H2)と関連しており、クリーチャの量がさらに少ないS1では、被験者2と被験者3の同時発話(H2H3)、被験者1と被験者2の同時発話(H1H2)の関連性が強い。

このような結果から、まずパーティ2の被験者においては、クリーチャの量の変化と発話者の交替に対応関係を見出すことができる。また、両方のパーティについては、クリーチャの量が少ない場合(S1, S2)の方がクリーチャの量が多い場合(S3)に比べ、被験者の同時発話と関連することが分かる。

以上から、両パーティについて、クリーチャの量の増減と同時発話という発話制御行動との間に関係性を見出すことができた。

### 3.4.5 まとめと考察

このような結果から、次の3つを確認した。一つは(1)クリーチャの量の増減と同時発話という発話制御行動を関連付けるパーティやその構成員にとって、クリーチャの量の人為的な変化は被験者の会話における発話行動に影響を与えるということを確認した。ここから、2つのパーティ間での会話の盛り上がり反応して2つのテーブル間を行き来するクリーチャの群れの動きが、2つのパーティにおける会話参加者の発話行動の調整に利用されていると考えることができる。

しかし、このような使用法をすべての組が行なっている訳ではない。実験では、1組(2つのパーティ)の被験者についてのみ確認することができた。これは、他の要因(例えば、会話の中での相手のジェスチャや話題の展開というNEXUS+のアニメーション以外の要因)によっても、このような発話の調整が行われていることを示している。

もう一つは(2)パーティ間の共同性を引き出す媒介物(メディアータ)としての働きである。半構造化インタビューの中で17名の被験者が回答したように、自分のグループのクリーチャの量が減少した時に発話が促されるなど、2つのパーティ間で「つなひき効果」のようなものが生じることを確認した。このように、NEXUS+が2つのパーティの間でのつながりを引き出すメディアータ(媒介物)になっている。これは、NEXUS+を囲んで2つのパーティが「一つのシステム」を作り上げることで、一体感のようなものが引き出されていると考えられる。つまり、クリーチャを引っ張り合うという活動を通じて、相互の間の秩序的なものを一緒に組織していると考えられる。

三つ目は(3)パーティ間での疎通性を引き出すことである。クリーチャの増減に関連して他方のパーティを気にかけるという行動を示す回答が12名の被験者から得られた。これは「一つのシステム」を介しての知覚行為といえる。それはどのような知覚行為なのかといえば、その場に対する能動的な関わりによって、その性質を探ったり、自分の身体を媒介として他者の様子をうかがっている、推測している。これは「なり込み」と呼ばれる現象である。その結果として、向こう側のパーティの被験者の性格が見えることがある。

これを双方のパーティで行っているとすれば、そのクリーチャの群れの動きを媒介にして「相互になり込む」状況を生みだしているといえるだろう。同時に、その相手との関わりの中で自分の性格などを見出している(グループとしての対人的な自己を獲得している)といえる。

NEXUS+ではそれぞれのパーティの盛り上がり反応してクリーチャの量が相対的に変化するという設定にしたことで、参加者は自分のテーブルのクリーチャの量から、もう一方のパーティの様子を推定することができると考えられる。こうして、自分たちのグループが感じていることを手掛かりに、相手のグループが感じていることを探ることができる。このような一連のシステムがグループとしての対人的な自己の獲得に繋がったと考えられる。また、会話参加者はこうして自分たちとの差異から相手のパーティの状況を推し測ろうとしていたのではないかと考えられる。

### 3.5 まとめと今後の展望

本研究では、スウィング・ルーム (swinging room) 中での人の揺動運動を調べた Lee らの Optic Flow Field に関する研究を手掛かりに、この Optic Flow Field のアイデアを社会的相互行為の場に展開し、その効果を確認することを試みた。TableTalkPlus を使用した先の実験では、会話参与者間の切り結びの結果として生じた Optic Flow を人為的に動かすことで、会話参与者間の会話における発話制御の調整行為に影響を与える効果が示唆されている。また、NEXUS+ を使用した実験では、2つのパーティの会話の結果として生じた Optic Flow を人為的に動かすことで、離れたパーティの会話参与者間の会話における発話調整行為に影響を与え得ることを明らかにした。

これらから、次のような意義や応用展開が考えられる。一つは (1) TableTalkPlus や NEXUS+ の枠組みを使用することで、Lee らの Optic Flow Field のアイデアを社会的相互行為の場に展開できるという点である。TableTalkPlus や NEXUS+ を介在させた多人数会話を分析することで、社会的相互行為の「場」を構成的に理解することができる。今後は、ジェスチャなどの NEXUS+ アニメーション以外の視覚的な要因を規制して実験を統制することで、より厳密なデータが得られると考えている。

もう一つは (2) ネットワークの向こう側に居るパーティとつながっている感覚を引き起こすことで、自分たちのグループ内の共同性を引き出すインタフェースとしての応用である。ここでは、会話における発話の量的な変化に着目して2つのパーティを比較することで、相手のパーティとの比較の中で自分たちのグループの性格などを見出す（グループとしての対人的な自己を獲得する）手段を提供した。これは「光学的流動」を間にして、「並ぶ関係」でのコミュニケーションを実現していると考えられる。

これを応用すれば、例えば、各パーティの会話の中心になっているキーワードや頻出する単語などを2つのパーティ間で比較することで、会話の質的な側面から、グループとしての対人的な自己を獲得するようなインタフェースを提供する

ことができる。

三つ目に (3)Social Network Service (SNS) などでは生じている社会的なインタラクションに対する一つの説明原理を提供できる可能性がある。Blog や Twitter などのソーシャルメディアへの自分の書き込みに対する他者からのコメントや返信を得ることで、ネットワーク上での他者との関わりにおける対人的な自己の獲得と調整を行っているという解釈である。このことについては、今回の実験では十分に明らかにするまでには至っていない。今後は特定の被験者の特異な行動に着目することで、より微視的な観点から分析を進めたいと考えている。

## 第4章 説得性の社会的構成に着目したメディア技術—タイプIII

### 4.1 はじめに

Human-Robot Interaction の研究分野における最も大きな話題の一つとして、一対一でのコミュニケーションを円滑にするターンテイキング構造を構築するために直接的なリソース（会話の状態やノンバーバルコミュニケーションにおける振る舞いの変化）をいかに活用するのかという議論がある [81, 45, 42]。多人数での会話を考慮した場合、ターンテイキングに関わる直接的なリソースを利用することは、会話参与者にとっても非常に複雑なものとなる。例えば、会話の状態や会話参与者の振る舞いのバリエーション、そして、会話における参与役割などが、会話参与者の増加に伴いより複雑になる [9, 27, 17]。このようなりソースは多人数での動的なインタラクションにおける会話参与の役割を決定するために柔軟に活用されるべきである。

フォーマル/インフォーマルな会話に関わらず、ある会話参与者だけが会話を独占し、一方で、他の参加者たちは会話に参加できないという事態を私たちは経験する。このような出来事が生じる場面では、会話参加への意欲が削がれている、会話参加者間で非生産的な会話が出来上がっている可能性がある。このような状況において、現在の会話の状態を認知的なアーティファクトを用いた記号的な表象によって、会話参与者に呈示することができるならば、会話への参加意識を促したり、より生産的な会話を引き出すような、参加者間の円滑なターンテイキング構造を確立するファシリテーターとして働く可能性がある。

認知的なアーティファクトは、表象的な機能として、情報を保存、表示、ある

いは、操作するためにデザインされた一種の人工的な装置である [61]。ここで表象的な機能は、人と人のインタラクションの特性、人工物がタスクに与える効果とその使用、人工物に対するユーザの視点、ユーザや専門家の技能によって変わる。人工物の表象（文脈に依存した機能）は外的表象と内的表象の2つに分類できる。理論的にいえば外的表象とは、表象を永続的に保存できるような記憶補助（例えば用紙等）にあたる。人工物は（何らかのインタラクティブメディアを介して）記号的な表象を提供することができる。このとき、人工物の内的表象は装置の内部で連動している。人工物の内的表象は、ユーザに向けて表現され、ユーザが使用する構造化された情報を記号化する一種のインタフェースとして捉えることができる [60, 14]。

ターンテイキング構造とは円滑で適切に発話を交替するために、会話参加者たちの役割を統合し、コミュニケーションの進行に際してペースを整えたり、適切なやり方から逸脱していないのかをチェックする、多様な機能を果たすコミュニケーションの構造の一つとして考えることができる [7, 15]。会話においてターンの交替を行うために、会話参加者は独自のルールセットを持っている。大抵の場合、それは会話を感知することで得られる。近年の研究 [57, 64] では、Goffman の「フットイング (footing[26])」という概念を用いて、会話の状態を把握・分析する試みが行われている。このような研究は複数人の人間に対して（一対一で）発話を交替するロボットを構築するうえで無視することはできない。また、会話における役割の交替規則を調査するうえでも重要である。これまでの先行研究 [107] では、会話のリズムを変化させるような媒介物として、インタラクティブメディア上のクリーチャを会話参加者がどのように利用するのかを調査した。その結果、上述のインタラクティブメディア（クリーチャ）が介在することで、会話参加者間での発話のターン交替がより意識される、発話したいという気持ちを引き出していることが分かった。

ここでは、会話参加者の参与役割を交替するために多人数会話におけるユーザの行動をアシストするという認知的なアーティファクトに基づいた内的表象（インタフェース）を提案する。アーティファクトに基づくインタフェースは人工物の

表現するシステムをユーザーの特徴にマッチしたものに变化していく。ここで述べる人工物の表現するシステムの特徴は、会話への参加意識を得るために、ユーザ同士のつながりを引き出したり、参与役割の交替に対して適応したりする。本研究では、Sociable Spotlight と呼ぶ人工物（インタラクティブメディア）を用いて、会話の流れに合わせて变化する内的表象を生み出す。このようなプロセスと連携して、会話参与者間のターンテイキング構造を変容させることは新たな手法といえる。

## 4.2 The Sociable Spotlights



図 4.1: 複数人での多人数会話に介入する Sociable Spotlights

これまでのソーシャルロボット [6, 64] や人工物 [84, 86] は直接的なインタラクション（例えば、一対一でのコミュニケーションや協同的なタスク）を確立することを前提としてきた。このような研究とは対照的に、Sociable Spotlight は傍観者として存在し、会話参与者たちの役割を交替させるようなソーシャルメディアエー

タとして振る舞う（図 4.1）。このように、人との直接的なインタラクションを要求せずに、認知的なアーティファクトにおける内的表象として駆動することは、人とロボットのインタラクションにおいて新しい試みといえる。ソーシャルメディアとしての記号的な表象とターンテイキング構造とを、ユーザはどのように（リソースとして）利用するのかを理解することは、ネットワークで繋がれた遠隔地間の（音声対話のみの）対談に応用できる。また、いかなる遅延も生み出さずに会話に対する参加意識を高めるための新しいインタラクティブインタフェースを生み出す際に欠かせない。

#### 4.2.1 認知的なアーティファクトの内的表象

Sociable Spotlight は会話における傍観者としての働きを持っている。会話参加者らの参与役割を交替するために用いられるユーザの会話状態は傍観者によって表現できると考えている。ここでは、会話の状態を記号的に表現する（人工物の情報処理の構造として描写されるようなスポットライトの発光色や動作）ために動的に変化する会話の要素を考慮した。

このような認知的なアーティファクトの内的表象（人工物の振る舞いや発光色の変化という記号的な表象）はユーザが会話を感じ取るに従って、理解・認識することができる。これは人間の情報処理の構造として考えられる。このような人間の情報処理と人工物の処理構造とが組み合わさることで、会話参加者は現在の会話の流れや他者とのつながりが表現されていることに気づく。他者とのつながりはターン譲渡表示として会話参加者に用いられるようなターンテイキング構造を引き出す。このような認知的な人工物の内的表象によって、会話参加者の参与役割を交替する（ターン譲渡表示を行う）、個人間の繋がりを引き出し、参与役割の交替に適応することで、参加意識の向上を狙う（図 4.2(右)）。



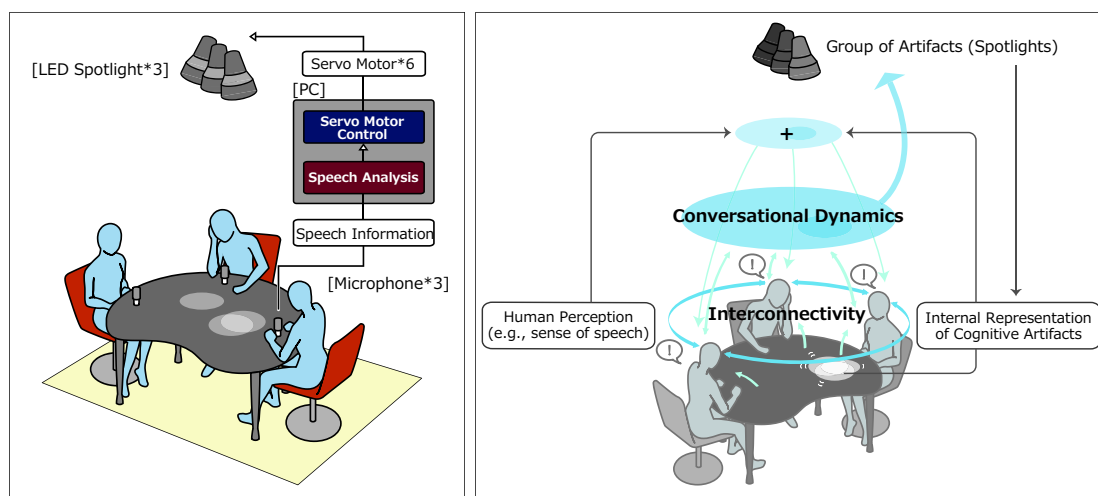


図 4.2: Sociable Spotlights のシステム構成 (左) と認知的なアーティファクトの内的表象を生み出すメカニズム (右)

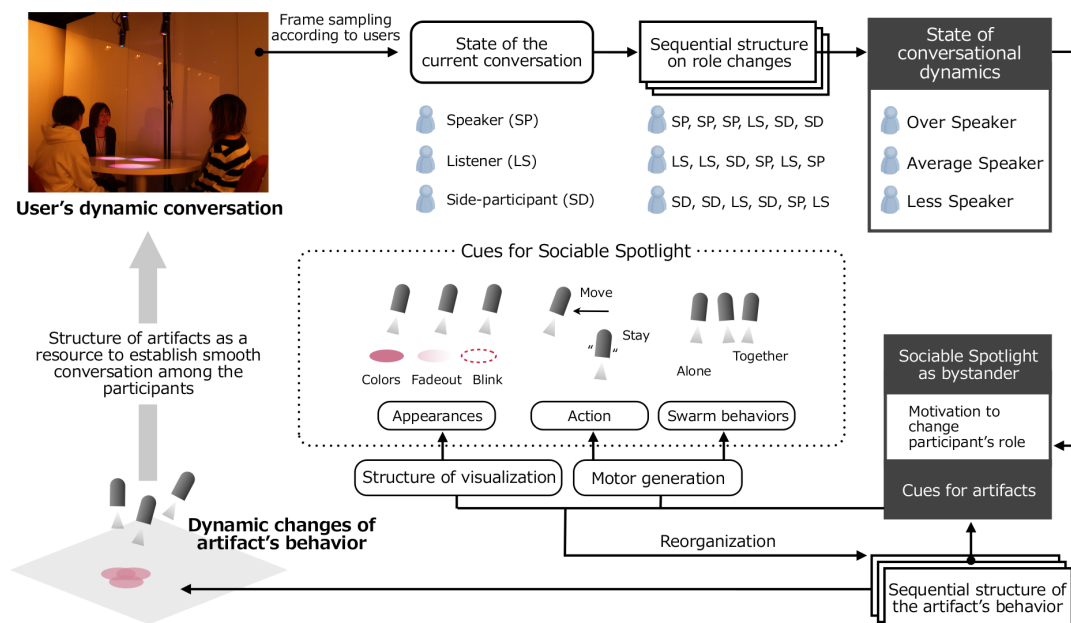


図 4.3: Sociable Spotlights のシステム構成と認知的なアーティファクトの内的表象を生み出すメカニズム

### 4.2.2 傍観者としての役割を果たすシステムの構成

図 4.2(左) と図 4.3 は 3 人での多人数会話に対して傍観者として参与する働きをもつ認知的なアーティファクトの構造を示している。現在の構成では、各マイクロフォンから得られる音声のパワー包絡に基づいて、会話参加者の発話交替を認識するために、3つのマイクロフォンを用いている。それぞれの会話参加者の会話の状態（発話者、聞き手、傍参加者）はしきい値を超える発話の音量（パワー）が得られたかどうかで決まる。インタラクションの初期では、10 フレームのデータのみを用いて、発話者や聞き手、傍参加者などの頻度を計算する。どの会話参加者が聞き手に回りがちで、どの会話参加者が話し手として多く会話に参加しているのかが分かる。その後、会話参加者の参与役割の履歴に従って、全体の会話の状態（話し手に回ることのない参加者、標準的な参加者、聞き手に回ることのない参加者）を把握する。ある会話参加者は何フレームの間、継続的に話し手に回っていたのか、どの会話参加者が何フレームの間、聞き手に徹していたのかなど、参与役割の交替に関わる連続的な構造に従って、本システムは会話の流れや状態を見積もる。そして、一人の会話参加者だけが聞き手に周りがちになり、他の 2 人の会話参加者だけでターンを交替しているような状況（とくに一人の会話参加者だけがターンを独占しているような状況が多い）では、本システムはその会話に参加できない会話参加者の方をライトアップして照らす。

会話の流れに従って、本システムは Sociable Spotlight の適切な振り舞い（スポットライトの動作や発光色の変化）を生成する。3つのスポットライト（発光部には Light Emitting Diodes（以下、LEDs と略記）を用いている）が実装されており、それぞれのスポットライトはサーボモータに付属している。それぞれのスポットライト、および、サーボモータは別々に動作させる（LEDs の発光色を別々に変化させる、別々に移動する）ことができる。これがターン譲渡表示となって会話参加者たちが利用できる。ソフトウェアは Max5.0 & Processing 上で駆動し、現在の会話の流れに合わせてスポットライトを動作（複数のスポットライトを同時/別々に移動・発光）させている。

3つの人工物（スポットライト）は参与役割を交替するために、会話の流れに合

わせて、群れのような動きを展開する。まず、3つの人工物は会話参加者の前を白色のスポットライトで照らす。つぎに、会話の状態に従って、一つひとつのスポットライトの発光色を変化させる。また、スポットライトの照射角度もサーボモータによって変化する。(詳しい動作原理を図 4.4 に示す)。そして、現在のターン譲渡表示が参与役割の交替に対して効果が無いものであれば(システムがそのように判断したら)、参与役割の交替が引き出されるまで、より効果的な別のターン譲渡表示(3つのスポットライトの発光色や照射角度の組み合わせ)を再構築する。

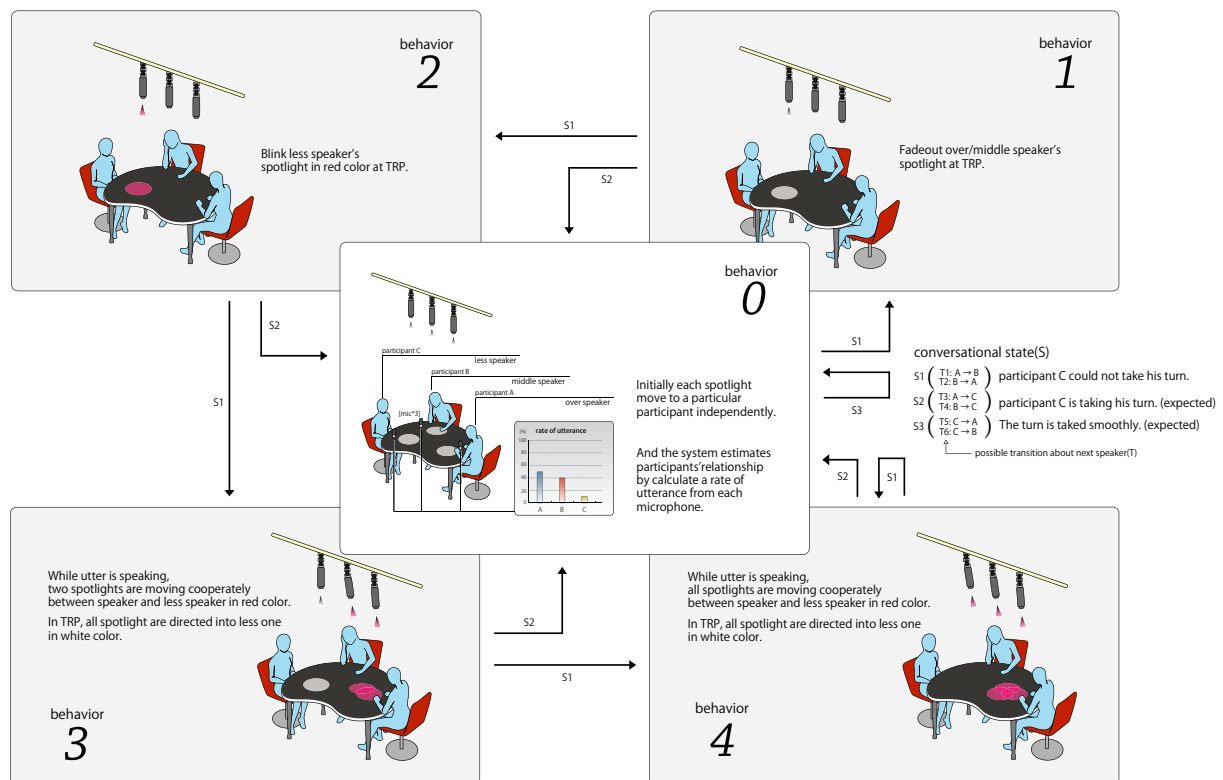


図 4.4: 参与役割を交替する目的で会話状態に基づき生成される人工物の振る舞い (図中では behavior(0) → SSB0, behavior(1) → SSB1, behavior(2) → SSB2, behavior(3) → SSB3, and behavior(4) → SSB4 と略記)

### 4.3 評価実験

スポットライトの振る舞いや記号的な表象が参与役割の交替にどのように利用されているのかを調査した。また、記号的な表象をリソースとして、会話参加者間のつながりがどのように確立されるのかも調べた。これはどのような認知的なアーティファクトの内的表象によって、会話参加者のつながりによってコミュニケーションを円滑にする、会話参加に対する意識を高めるようなターンテイキングの構造が引きだされるのかを把握する上で重要である。このような目的のために、会話の内容や話題を認識することなく、被験者の発話のパワー包絡（頻度）のみの情報をシステムに与えた。

実験は2つのセッションから構成される。(1) まず、それぞれのグループのメンバー同士で（スポーツ中継の後半戦の）5分程度のビデオを一緒に観戦する。(2) 1のビデオに関して、認知的なアーティファクトの内的表象を介しながら、5分間で自分たちの感想を述べる。このような会話の流れで、提案するインタラクティブアーティファクトは被験者の会話参加への意識を高めたり、円滑な発話の交替を促す。セッション1では、実験室内での自然な会話を引き出すための話題作りに欠かせない。

12名（3人1組、計4組）の被験者がこれまでに実験に参加している。4種類の人工物の振る舞いの効果を確認した。

(0) 聞き手に回りがちな参加者は次の発話のターンを得ることができない (1) 0の状態であれば、Turn Relevance Place（以下、TRPと略記）において、発話を独占する参加者と中間的な参加者の前を照らす白色のスポットライトをフェードアウトする（SSB1と略記）(2) 1を行い、0の状態であれば、TRPにおいて、聞き手に回りがちな参加者の前を赤色のスポットライトでチカチカと強調する（SSB2と略記）(3) 2を行い、0の状態であれば、2つのスポットライトが発話を独占する参加者と聞き手に回りがちな参加者の2人を赤色のスポットライトで交互に照らし、その後、TRPにおいて3つのスポットライトが聞き手に回りがちな参加者の方を白色のスポットライトで照らす（SSB3と略記）(4) 3を行い、0の状態

あれば、3つのスポットライトが発話を独占する参加者と聞き手に回りがちな参加者の2人を赤色のスポットライトで交互に照らし、その後、TRPにおいて3つのスポットライトが聞き手に回りがちな参加者の方を白色のスポットライトで照らす（SSB4と略記）

以上の遷移を図4.4に図示する。会話の流れに合わせて、システムは発話を独占する参加者と聞き手に回りがちな参加者、中間的な参加者を選別し、それぞれの会話状態に合わせてスポットライトの動作/明度/発光色を計算し、Sociable Spotlightsの動作を生成している。

## 4.4 結果

実験の詳細な分析の前に、スポットライトの振る舞いと被験者の参与役割の交替の間の関係を調べるのが重要である。

結果は、被験者が認知的なアーティファクトの内的表象を参与役割を交替するための一種のリソースとして、その会話の流れの中でどのように利用したのかを示している。スポットライトがどのような振る舞いを行ったとき、被験者の発話交替が生じたのかをカウントした。カイ二乗検定を用いて、その関係性を調べた。その結果、スポットライトの振る舞いと被験者の参与役割の交替の間に有意な差があることが認められた ( $\chi^2=11.7$ , d.f.=3,  $p=0.008$ ,  $p<0.01$ )。以上のことから、会話の流れの中での人工物の振る舞い（例えば、発光色/照射角度の組み合わせの効果）が参与役割の交替に用いられていることが伺える。

もう一つのエビデンスを表6.1にまとめた。表4.1はスポットライトの各動作（連携）に対して、被験者の参与役割がどの程度生じたのか（参与役割交替の成功率）を示している。さらに、“System”の列は図4.4で述べたスポットライトの振る舞いを示している。“Numbers”の列は参与役割を交替する前の役割（“Initial user’s role”）からターン交替に伴って、どのような役割に移行したのかを回数で示している。この回数に従えば、人工物の振る舞いの変化（遷移）するに連れて、参与役割の交替する頻度（割合）に顕著な差が生じていることが分かる（“User’s

表 4.1: スポットライトの各動作に対する被験者の参与役割交替の成功率（図中では、発話を独占する参加者 → Over, 中間的な参加者 → Average, 聞き手に回りがちな参加者 → Less と略記

System		User's feedback			
Structure of artifact		Initial user's role	Turn-taking	Numbers	Successful rate
Behaviors	Structure of the visualization				
SSB1	SSB0 → SSB1	Over/Average	Less	24	40.68%
		Less	Over/Average	35	59.32%
SSB2	SSB0 → SSB1 → SSB2	Over/Average	Less	10	71.43%
		Less	Over/Average	4	28.57%
SSB3	SSB0 → SSB1 → SSB2 → SSB3	Over/Average	Less	2	100.00%
		Less	Over/Average	0	0.00%
SSB4	SSB0 → SSB1 → SSB2 → SSB3 → SSB4	Over/Average	Less	9	60.00%
		Less	Over/Average	6	40.00%

feedback” 列)。

#### 4.4.1 記号的表象を介した被験者間のつながり

日常的な会話は発話者からの問いかけと、聞き手からの応答によって成立していることが一般に知られている。このことから、その参加者はいま会話にどの程度、従事しているのかを判断するためには、参与者同士のつながりを調べることである程度、把握することができる。このようなことから、参加者同士のつながりを形成する上で、認知的なアーティファクトの内的表象を被験者たちはどのように利用するのかを探った。認知的なアーティファクトの内的表象と被験者の振る舞い（発話者としての参加（以下、SR と略記）、聞き手としての参加（以下、LS と略記））との関連性を調べた。ここでは、代表的な2つのグループを取り上げて説明する。他のグループについても、同様の傾向が得られている。

コレスポネンス分析による描写的/探索的な分析は被験者間のつながりを取り出すために用いている。コレスポネンス分析は表データを視覚的に簡素に分析するために開発されたもので、群（列と行）間の関連性を測る際に用いられる。関連線の高い要素同士はより近い位置にプロットされる。頻度データ（例えば、スポットライトの動作に対する被験者の聞き手/話し手として参加した回数）をプロットすれば、それぞれの関係を互いの距離という情報から探ることができる。

コレスポネンス分析の結果を図に示す。グループ1のそれぞれの被験者の会

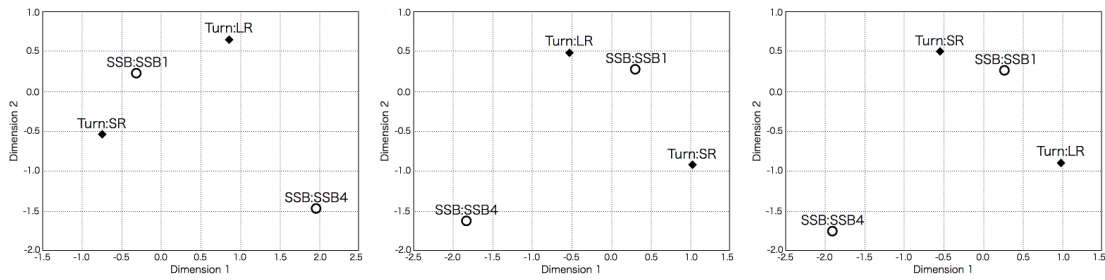


図 4.5: 発話を独占する参加者（左），中間的な参加者（中央），聞き手に回りがちな参加者（右）のコレスポネンス分析結果

話の状態を示している。図 4.5(右) は聞き手に回りがちな参加者の結果である。図 4.5(中央) は中間的な参加者の結果である。図 4.5(左) は発話を独占する参加者の結果である。結果から、発話を独占する参加者や中間的な参加者では、聞き手としての参加 (LS) と SSB1 というスポットライトの振る舞いとが互いに近い位置にプロットされる。一方、聞き手に回りがちな参加者では、話し手としての参加 (SR) と SSB1 というスポットライトの振る舞いとが互いに近い位置にプロットされる。以上のことから、聞き手に回りがちな参加者が話し手として参加 (SR) した際 (図 4.5(右)) に、発話を独占する参加者と中間的な参加者は、聞き手として参加 (LS) している (図 4.5(左)) ことが分かる。

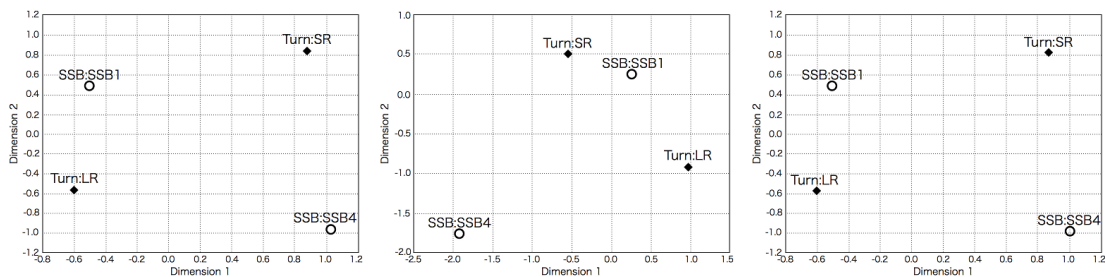


図 4.6: 図中のコレスポネンス分析結果は発話を独占する参加者（左）と聞き手に回りがちな参加者（中央）と中間的な参加者（右）と人工物の内的表象の関係を示している

グループ 2 の会話状態において、3 人の被験者で個人間のつながりが現れたのは非常に興味深い (図 4.6(左), 図 4.6(中央), 図 4.6(右)) SSB1 で発話に消極的な態

度を示す参加者が話し手としての参加を行った（SR と SSB1 が近い距離にある）ことが分かる（図 4.6(中央)）。このとき、発話を独占する参加者（図 4.6(左)）と中間的な参加者（図 4.6(右)）は聞き手に回っている（LS と SSB1 が近い位置にある）。このような結果から、現在の会話状態を感知して、認知的なアーティファクトの内的表象を用いることで、被験者たちは互いに連携していることが伺える。このような連携によって、発話の交替を円滑に進めるようなターンテイキング構造が引きだされている。

この結果から、認知的なアーティファクトの内的表象の効果が明らかになった。はじめ、被験者らは認知的なアーティファクトの記号的な表象に対する知識がない。しかし、会話の感知（人間の情報処理の構造）と認知的なアーティファクトによる記号化とが関連づけられた後では、被験者たちはターン譲渡表示を用いて他の会話参与者とのつながりを確立することができた。各グループの独自の手法に従って、認知的なアーティファクトの内的表象は会話参与者間のターンテイキング構造を構築するために用いられていることが分かった。

## 4.5 まとめ

本研究では、多人数会話において社会的な人工物（Sociable Spotlights）として駆動するユニークな手法を提案した。ただ会話を傍観するだけの認知的なアーティファクト（Sociable Spotlight）にも関わらず、現在の会話の状態が記号的に表現されることで、会話参与者たちは現在の会話状態を考慮して、発話を行っていることが分かった。会話参与者たちは会話の感知（情報処理の構造）と認知的なアーティファクトの内的表象とを関連付けて、自分なりのターンテイキング構造を構築し、会話参与者間でのつながりを確立していた。このような会話参与者間でのつながりは、認知的なアーティファクトによってインタラクティブに生まれた「会話のダイナミクス」を記号的に表象したことで現れたものである。以上のような、他者間でのつながりを生み出すプロセスは会話参与者たちのターンテイキング構造によって反映・誘発されたものである。総論すると、ターンテイキング構造を



形づくる重要な要素として利用され、会話参加者間のつながりを引き出すことで、会話に対する参加意識の向上を促すという認知的なアーティファクトの内的表象の効果を実験結果によって示すことができた。



# 第5章 説得性の社会的構成に着目したメディア技術のHAIへの応用例

## 5.1 はじめに

風による影響で、道端に落ちていた白いビニール袋がこちらに向かって自然と動きはじめた。遠くから眺めていると、単なるビニール袋が「誰かに構って欲しい」という意図をもった小動物のようにも思えてくる。このような思い込みによく似た現象は、Heider & Simmelが作成したアニメーションの中にもみられる。

Heiderらは、大小の異なる三角形と小さな円の関係性をデザインし、これをアニメーションとして作成した<sup>1</sup>。被験者に与えられた課題は、この動きまわる一つひとつの図形の特徴を説明するというものである。被験者は生物とは全く異なる見た目の図形の動きに対しても、「追いかける」「逃げる」「戦う」などの意図に関連する言葉を用いて、その図形の特徴を説明したという [33]。

これと関連する研究として、白色と黒色の二つの円を用いたBassiliらの実験 [4]がある。Bassiliらは、二つの円の運動変化のタイミングが随伴的であっても、二つの円の運動方向がランダムである場合では、被験者が図形の動きに意図を見いだすことはないこと。しかし、同様の時間条件が他方の運動方向に追従するという動きと組み合わせることで、単なる図形の動きに意図を見いだすということを明らかにした。

---

<sup>1</sup>Heider, F., and Simmel, M. A.: Heider-Simmel Demonstration, [www.youtube.com/watch?v=sZBKer6PMtM](http://www.youtube.com/watch?v=sZBKer6PMtM) (1943).

このような研究成果が与える新たな視点は、見かけや形状という図形自体の特徴からではなく、図形同士の因果的および空間的な関わりの中で、対象の意図が構成されている点である。ただし、Heider や Bassili らの実験では、仮想世界 (CG) におけるインタラクションに議論の焦点が絞られていた。本研究では、これまで CG の中で実現していた前述のアプローチを、人とのインタラクション (=社会的な関わり) に展開 (= Sociable) する。これにより、Human-Agent Interaction の枠組みにおいて、モノのエージェンシーを構成する新たなアプローチが実現する。モノのエージェンシーを構成する試みには、次のような意義がある。

あるモノにエージェンシーが帰属されるとき、そのモノは自分の意志で動いているという印象を人に与え [41, 124]、人々の行動や社交空間を変容させるような働きをもつ [31]。人とモノ (ロボットやソフトウェアエージェントなど) とのリアリティを伴ったコミュニケーションの実現には、そのモノのエージェンシーを人が感じ取れることが欠かせない。これまでに、モノのエージェンシーを形づくる試みは様々なアプローチから行われてきた。

例えば、ロボットの姿形を動物のような愛らしい形態にする、ロボットの外見を人間に酷似させる、動物や人間の振る舞いをモデル化しロボットに実装する、適応学習や発達のメカニズムを用いるなどの方法が挙げられる。Human-Agent Interaction における一つの課題として、モノへのエージェンシー帰属に関する議論は現在も進められている [102, 109]。この他にも、人の意識や行動をさりげなく変えるためのモノ・コトを体系的に理解する仕掛学 [98] の研究アプローチも、エージェンシーによる人の行動変容を議論するうえで無視することはできない。また、参加者のエージェンシーを顕在化するようなインタラクティブメディア [63, 94] についても、人のエージェンシーが形づくられるメカニズムを解明するうえで重要である。

このように様々なアプローチが検討されている中で、本研究は Heider や Bassili らの着眼点を社会的な関わりに展開するという立場 [104] である。この立場に従えば、従来の Human-Agent Interaction 研究が前提としてきた「モノのエージェンシーはそのモノ自体に備わる」という考え方から脱却し、新たなパースペクティブを提供するものと思われる。モノのエージェンシーは人と人との社会的な関わりの中で生

み出され、人と人の社会的な関わりを方向づけるアーティファクト (artifact[105]) である。この観点が本研究を進める上での基本的なアイデアであり、この観点に基づいてモノのエージェンシーについて議論することが本研究の目的である。

そこで本研究では (1) 社会的な関係の中で形づくられるエージェンシーに着眼する実験システム (Sociable Spotlight) を構築し、(2) この Sociable Spotlight におけるエージェンシーの構成と、人からの社会的な関わりとが密接に関係していることを、体系的な分析を通じて明らかにした。本論文では、はじめに実験システムのデザインと構成について整理する (5.2 節)。つぎに、Sociable Spotlight のエージェンシーが人と人との社会的な関わりの中で生じていることを論じ (5.3 節)、人からのどのような関わりが Sociable Spotlight におけるエージェンシーの構成に関連していたのかを質的および量的の横断的な分析を通じて理解することで、5.3 節の議論を補強する (5.4 節)。最後に本研究の明らかにしたことを整理し、今後の課題や展望について述べる (5.5 節)。

## 5.2 実験システムのデザインと構成

実験に使用したシステムのデザインと、そのハードウェア構成や内部での処理について詳しく述べる。

### 5.2.1 デザインのポイント

Heider らの研究や Bassili らの実験結果に倣い、観察に使用するシステムを社会的なインタラクションに拡張できるように設計した。そのポイントは次の 3 つにまとめられる。

#### テーブルというアーティファクト

テーブルを囲んで人々は日常的に雑談をし、娯楽を楽しむ。テーブルという家具自体が人と人の社会的な関わりを予定し、人と人の社会的なインタラクション

を引きだすアーティファクトとしての働きを持っている。このようなテーブルをシステムの一部として用いることで、エージェンシーの形成に関連深い社会的なインタラクションを引きだす。

### 随伴的な関係性

Bassili らの研究では、物体と物体の間の随伴的な関係性に着目した。ここでは、これを社会的なインタラクションに展開し、対象物（モノ）と社会的な他者（ヒト）の間において随伴的な関係を結ぶ。ここでは周囲にいる複数人の参与者（ヒト）が発する音声のパワー（音量）変化のタイミングに随伴して対象物を移動することで、モノとユーザの間で随伴的な関係を構築した（5.2.3で詳しい動作を述べる）。

### ミニマルデザイン

Heider らのアニメーションと同様に、ここでは単純な円形のアニメーション（後述の Sociable Spotlight）を対象物として用いた。このようなミニマルデザイン [119] を用いることで Sociable Spotlight に対する周囲の参与者からの積極的な解釈や関与を引きだす。例えば、参与者によつては Sociable Spotlight が何らかの志向に従って動いているものと解釈をし、生き物のように扱う、その動き方に生物学的な意味合いを与える、などの行為が引き出される。このような人からの社会的な関わりがあつて、Sociable Spotlight のエージェンシーは形づくられると考えている。

## 5.2.2 システム構成

5.2.1の3つのポイントを満たすシステムを構築するために、システムの概観を図5.1のようにした。このシステムは複数人での多人数インタラクションを想定したインタラクティブメディアである。

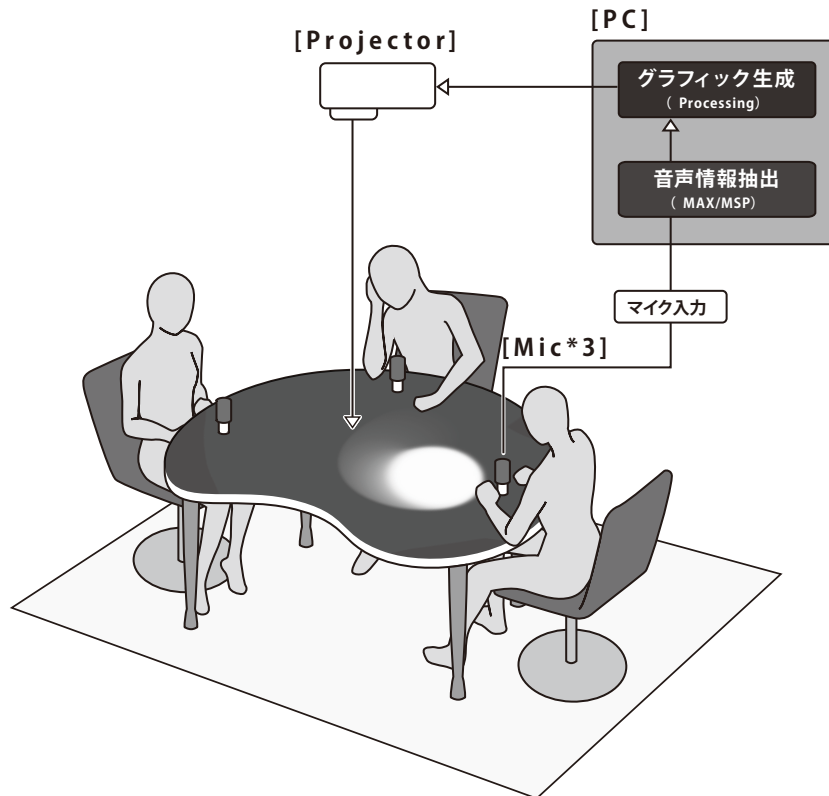


図 5.1: システムの概観 (参加者の取り囲むテーブル上に図 5.2 のアニメーションを投影)

### ハードウェア

図 5.1 に示すように、ハードウェアは周囲の環境音を收音するワイヤレスマイクが 3 機 (テーブルの下に正三角形の頂点に位置する角度で装着)、マイクレシーバが 3 機、これらを PC に接続するための USB オーディオインタフェースが 1 台、アニメーションの生成及び制御処理用の PC、アニメーションを投影するための液晶プロジェクタと映像の投影先となるテーブルからなる。

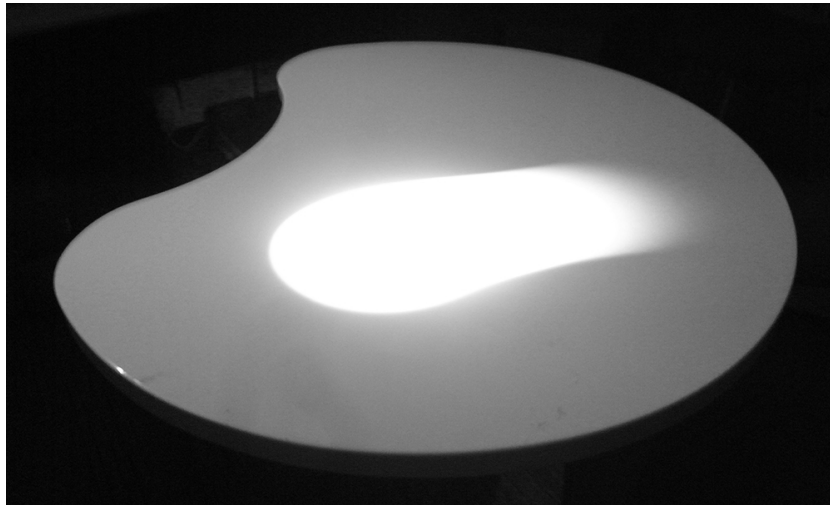


図 5.2: 環境音の変化のタイミングに合わせて随伴的に揺れ動く Sociable Spotlight アニメーションを映したテーブル

### ソフトウェア

PC内のソフトウェアはMAX/MSPを用いてそれぞれのマイクロフォンから得られる音声のパワー包絡を解析し、この情報に従ってProcessingによりアニメーションを生成する。生成されたアニメーションはPCからプロジェクタへ出力され、テーブル上に投影される。

ここでは上記システムに Sociable Spotlight と呼ぶアニメーションを実装した(図 5.2)。直径 25cm の白い円型の図形がテーブル(横 110cm × 奥行 95cm × 高さ 72cm)に投影され、最も接近する場所でテーブルの淵から約 10cm 内側を周囲の環境音の変化に合わせて 0cm/s から約 20cm/s の間の速度で滑らかに移動する<sup>2</sup>。

<sup>2</sup>動作例：<http://youtu.be/JAI5FiHe94k>



### 5.2.3 Sociable Spotlight の移動関数

Sociable Spotlight の詳しい動作を以下に示す。ここで  $x, y$  を Sociable Spotlight のテーブル上の現在位置 ( $X, Y$ ) とすると

$$x = x + V_x \quad (5.1)$$

$$y = y + V_y \quad (5.2)$$

で示すことができる。現在位置 ( $X, Y$ ) は 1 秒間に 60 回更新<sup>3</sup>される。  $V_x, V_y$  は Sociable Spotlight の速度で

$$V_x = \mu(V_x + a_x) \quad (5.3)$$

$$V_y = \mu(V_y + a_y) \quad (5.4)$$

である。  $\mu$  を摩擦係数 ( $= 0.95$ ) とする。  $a_x, a_y$  は Sociable Spotlight の加速度で

$$a_x = F_x/m \quad (5.5)$$

$$a_y = F_y/m \quad (5.6)$$

である。  $m$  を質量 ( $= 13.0$ ) とする。  $F_x, F_y$  は弾性力で

$$F_x = k(X_g - x) \quad (5.7)$$

$$F_y = k(Y_g - y) \quad (5.8)$$

である。  $k$  をバネの剛性 ( $= 0.05$ ) とする。  $X_g, Y_g$  は目標位置で

$$X_g = X \text{ 座標 (MIC}_n \text{ | テーブルの中央)} \quad (5.9)$$

$$Y_g = Y \text{ 座標 (MIC}_n \text{ | テーブルの中央)} \quad (5.10)$$

である。式 5.9, 式 5.10 はそれぞれ、パワー (音量) の最も大きいマイクロフォンの座標位置 ( $X, Y$ ) を返す関数である。ここで  $n$  をパワー (音量) の最も大きいマイクロフォンの固有番号とし  $0 < n \leq 3$  とする。また、Sociable Spotlight がテーブルの淵に到達した場合はテーブルの中央の座標が目標位置となる。

式 5.1~式 5.10 により、次の 4 つの基本動作が生まれる。

<sup>3</sup>滑らかな移動のために描画の際は軌跡を表示している。

**動作 1**

テーブル下の3つのマイクロフォンの内、最も大きい音量が得られたマイクロフォンの方向に Sociable Spotlight は移動する。

**動作 2**

加速度（式 5.5, 式 5.6）の増加により、Sociable Spotlight がテーブルの淵からはみ出しそうな（式 5.1, 式 5.2 で示される座標がテーブルの淵から約 10cm 外側に到達した）場合は、約 1 秒間、逆方向（テーブルの中央）に移動する。

**動作 3**

動作 1～動作 2 を繰り返しながら（式 5.9, 式 5.10）、徐々に Sociable Spotlight の速度（式 5.3, 式 5.4）は減速する。

**動作 4**

最大音量の得られるマイクロフォンが変化するたびに、動作 1～動作 4 が生じる。

以上のように、一つひとつの動作はシンプルではあるが、周囲の環境音の変化に従って移動することで、Sociable Spotlight の動作は不規則的なものになる。従って、Sociable Spotlight の動作原理を理解することは困難になり、加えて、5.2.1 の「ミニマルデザイン」によって、参加者は参加者なりの解釈を行うことが必要となる。

## 5.3 観察 I—公共施設でのインタラクション

### 5.3.1 観察の概要

Sociable Spotlight のエージェンシーが人からの社会的な関わりの中で構成されていることを明らかにすることが本観察の目的である。そのために (1) Sociable Spotlight に関わろうと試みる参加者に着目し、1 の場合に (2) Sociable Spotlight



図5.3: 左：手づかみを行うグループ（祖父—孫，他4組），中央：引っ張り合うグループ（幼児—幼児），右：シルエットを利用するグループ（父親—娘，他1組）を媒介にして，複数人の参加者を巻き込んだ社会的なインタラクションは生じるのか。2の場合に（3）Sociable Spotlightのエージェンシーはその社会的なインタラクションの中でどのように生じるのかを観察した。

### 5.3.2 観察手法

#### 観察のフィールドおよび被験者

様々なタイプのインタラクションを引き出すために，インタラクションの目的を定めない状態で，オープンスペースとしての公共施設の中に Sociable Spotlight を介在させた。様々な業種，幅広い年齢層の人々が往来する豊橋市の交流施設「子ども未来館ココニコ」を観察のフィールドとした。中でも Sociable Spotlight に初めて関わる瞬間から立ち去る瞬間までの来場者の振る舞いを対象としている。なお，来場者は Sociable Spotlight の動作原理およびコンセプトを一切知らされていない。来場者は公共施設内を自由に行き来している。

#### 録画環境

Sociable Spotlight とインタラクションする来場者たちの様子をビデオカメラで撮影した（2日間 2時間 36分収録）。交流施設内での撮影については Sociable Spotlight とインタラクションする来場者の様子をビデオカメラで収録しているこ

とを公示する張り紙を周囲に貼付する等して、プライバシーに配慮した。また、実験実施者の振る舞いが来場者たちの振る舞いに影響を与えることを最小限に抑えるために来場者の遠方から撮影した。なお、本実験は施設スタッフの協力のもとによる。

### 5.3.3 観察結果

#### 抽出手順

表 5.1 の手順に従い、収録したビデオデータから本観察で取り上げたい (5.3.1 の「観察の概要」の中で言及した) 3つの現象を抽出した。

#### 抽出結果

表 5.2 に示す通り、全体の収録データから A-1 に該当する来場者を 19 名確認した。この 19 名のうち A-2 に該当しない来場者を 6 名確認した。この 6 名の内訳として、すぐに立ち去る来場者が 2 名、視線を向けたまま何もしない来場者が 4 名であった。残りの 13 名は A-2 に該当した。例えば、手で触れるというもの、そのあと話しかけるといもの、影をつくるなどのインタラクションが観察された。この 13 名のうち A-3 に該当しない来場者を 4 名確認した。Sociable Spotlight と個人的なインタラクションをした後、最も短いケースで約 1 秒、最も長いケースで約 22 秒でインタラクションを終える。また、この A-3 に該当しない 4 名の平均インタラクション継続時間は約 9 秒であった。

残りの 9 名は A-3 に該当した。その内訳は次のとおりである。一つは、Sociable Spotlight を生き物に見たてて手で押さえつける、触れるなどして一緒に捕まえようと対話する来場者を 5 組確認した (図 5.3 左)。二つ目に「何これ」という来場者の対話をきっかけに、Sociable Spotlight を引っ張り合おうと互いに試みて、その生物性を探ろうとする来場者を 1 組確認した (図 5.3 中央)。三つ目に、Sociable Spotlight の影を利用したシルエット遊びをしながら対話を行う来場者を 2 組確認

表 5.1: 抽出の手順

手順	内容
A-1	Sociable Spotlight に 1 秒以上注視した来場者を抽出
A-2	A-1 の来場者で Sociable Spotlight と対話する、接触するなどのインタラクションを行う来場者を抽出
A-3	A-2 の来場者が Sociable Spotlight を媒介に別の来場者と対話や共同行為を行った場合、この場면을抽出

表 5.2: 抽出の結果

項目	該当	例外	該当例 (関与の方法)
A-1	19 名	-	1 秒以上の注視
A-2	13 名	6 名	接触, 対話, 影作り
A-3	9 名	4 名	手掴み, 引張り, 影遊び
平均時間	約 41 秒	約 9 秒	

した (図 5.3 右)。四つ目に、Sociable Spotlight の動作に合わせて、交互に Sociable Spotlight を手で触れるという行為を何度か繰り返した後に、「わっ」という発話を Sociable Spotlight に投げかけて Sociable Spotlight からの反応を探ろうと試みる来場者を 1 組確認した。

これら A-3 に該当する 9 組の Sociable Spotlight とのインタラクションの継続時間を計ると、最も短いケースで約 9 秒、最も長いケースで約 98 秒であった。インタラクションの平均継続時間は約 41 秒である。

#### 5.3.4 まとめ

以上を考慮すると次の 3 つが伺える。一つは、A-1、A-2 に該当する来場者が 13 名いることから、Sociable Spotlight に何らかの関わりが引き出されている。二つ

目に、A-1、A-2、A-3に該当する来場者が9名いることから、Sociable Spotlightを媒介にした社会的なインタラクションが生じている。ここではSociable Spotlightを生き物に見たてて一緒に遊ぼうと試みる来場者やSociable Spotlightからの反応を観察しながら、別の来場者との社会的なインタラクションを通じてその生物性を探ろうとする来場者がいた。三つ目に、Sociable Spotlightに対して1名で関わる場合（A-1、A-2に該当しA-3に該当しない4名）と、複数名で関わる場合（A-1、A-2、A-3に該当する9名）ではインタラクションの平均継続時間が顕著に異なっている（前者は約9秒であるのに対し、後者は約41秒であった）。複数名の参加者が互いに対話を交わすという状況の中でSociable Spotlightを生き物に見たてる共同行為やその生物性を探ろうとする共同行為が比較的長く引き出されている。ここから、複数名の参加者が互いに対話を交わすという状況はSociable Spotlightのエージェンシーを構成するうえで無視することのできない要素であることが分かる。

このような結果が得られた理由として、Sociable Spotlightの設計にミニマルデザインを用いたことが挙げられる。5.3.3の「抽出結果」に代表されるように、Sociable Spotlightの動作は子どもたちから「なにこれ？」という疑問を引き出していた。来場者はSociable Spotlightの動作原理を理解することが困難であり、自分なりの解釈や答えを探らなければならない状況にあったことが伺える。しかし、このような解釈の余地があったことで、Sociable Spotlightを生き物に見立てるというオリジナルな解釈が引きだされたものと考えられる。例えば、物珍しい玩具によって子どもたちの興味が惹き付けられていたというよりは、何かよく分からない物体とのコミュニケーションの試行錯誤にあつて、結果的にインタラクションの継続時間が長くなったと思われる。インタラクションの継続時間が長くなるという傾向が得られた背景には、このような要因があつたためといえる。

しかし、本観察は遠方からの撮影であつたため、具体的にどのような発話のやり取りの中で、来場者はSociable Spotlightのエージェンシーを構成しようと試みていたのかが定かではない。次章では、発話行動という社会的なインタラクションに着目することで、被験者のやり取りの細部に限定した分析を行い、5.3の議論を補強する。

## 5.4 観察 II—実験室環境における多人数会話

### 5.4.1 観察の概要

つぎに、エージェントの形成に有効と考えられる「互いに対話を交わすような状況」を実験室環境の中で再現した。ここでは多人数会話の場に Sociable Spotlight を介在させ、単なるアニメーションとしての Sociable Spotlight が (1) 話題としてその会話に取り込まれたときや (2) 発話者からの語りかけを支える「受け手としてのエージェント」を備えたとき、そして (3) 会話の流れの中で利用される価値が見出されなかった場合の被験者の発話行動の差異に着目した。このような発話行動の違いを手がかりに、どのような会話の流れで Sociable Spotlight のエージェントは形成されやすい/されにくいのかを探った。

つぎに、5.3 章での議論をサポートするエビデンスを得る目的で二つ目の観察 II を行った。ここでは、互いに対話を交わすという状況を実験室環境の中で再現し、5.3 章の観察 I では扱えなかった発話行動の細部に着目した分析を行う。

具体的には、多人数会話の場に Sociable Spotlight を介在させ、単なるアニメーションとしての Sociable Spotlight が (1) 話題としてその会話に取り込まれたときや (2) 発話者からの語りかけにより、発話の受け手としてのエージェントが形づくられたとき、対照的に (3) 会話の流れの中で利用される価値が見出されなかった場合の被験者の発話行動の差異に着目した。このような 3 つの発話行動の違いを手がかりに、具体的にどのような発話行動と Sociable Spotlight におけるエージェントの形成が関連しているのかを探った。

### 5.4.2 観察手法

#### 被験者

被験者として、Sociable Spotlight を見たことがなくシステムの動作原理に関する知識のない者を対象とした。互いに友人関係の 3 名を 1 グループとし、計 4 グループ (男性 6 名, 女性 6 名, 計 12 名, 平均年齢 20.7 歳 (20 歳–24 歳) の学内の

学生など) が本実験に参加した。

### 移動関数の調整

多人数会話に介在することを想定し, Sociable Spotlight の移動関数の式 5.9, 式 5.10 を

$$X'_g = X \text{ 座標 (参与者}_n \text{ | テーブルの中央)} \quad (5.11)$$

$$Y'_g = Y \text{ 座標 (参与者}_n \text{ | テーブルの中央)} \quad (5.12)$$

のように調整した。ここで  $n$  をグループ内で累積発話区間の最も短い被験者の固有番号とし  $0 < n \leq 3$  とする。また, 累積発話区間の算出できない (まだ一人も発話していない) 場合はテーブルの中央の座標が目標位置となる。また, 3つのマイクロフォンを一人ひとりの被験者の口元にくるように設置した。

具体的な動作例<sup>4</sup>を以下に示す。特定の会話参与者に Sociable Spotlight が近づく/遠ざかる状況は次のような場面である。

#### 近づく状況 I

特定の会話参与者 A の発話停止中に別の会話参与者 B が発話をしていたため, 先の会話参与者 A の発話継続時間が他の会話参与者 B よりも低下したと PC は判断し, Sociable Spotlight が先の会話参与者 A の方に近づく

#### 近づく状況 II

特定の会話参与者 A の発話中に別の会話参与者 B がより大きな音量で発話をしていたため, 先の会話参与者 A の発話継続時間が他の会話参与者 B よりも低下したと PC は判断し, Sociable Spotlight が先の会話参与者 A の方に近づく

#### 遠ざかる状況 I

特定の会話参与者 A の発話中に他の会話参与者 B が発話を停止していたた

<sup>4</sup>動作例: <http://youtu.be/SQM033-Eu-0>



め、先の会話参与者 A の発話継続時間が別の会話参与者 B よりも上昇したと PC は判断し、Sociable Spotlight が先の会話参与者 A から遠ざかる

### 遠ざかる状況 II

特定の会話参与者 A が他の会話参与者 B の発話中により大きな音量で発話をしていたため、先の会話参与者 A の発話継続時間が別の会話参与者 B よりも上昇したと PC は判断し、Sociable Spotlight が先の会話参与者 A から遠ざかる

このように、発話継続時間の相対的な関係に従って Sociable Spotlight が移動することで、一人ひとりの会話参与者からは不規則的に移動しているように見え、動作規則の理解を困難なものにした（会話参与者たちが自由に解釈できる余地を残した）。

### 実験課題

実験課題として、実験開始前に全員で一緒に視聴したビデオの内容に対する共同想起（joint remembering）対話 [16] を用いた。

### 手続き

被験者は実験開始前に「ビデオの内容を覚えてストーリーを正確に話す」という課題が伝えられた。また、被験者に対するプライバシーへの配慮、及び実験により生じる被験者への不利益及び危険性について説明し、最後に「被験者はこの実験に参加しない自由をもち、また実験に参加してもいつでもその同意を撤回することができる」という旨を伝えた。

実験の教示内容は次のとおりである。(1) ビデオ視聴ブースに移動し、5分程度の娯楽アニメ（Tom and Jerry<sup>5</sup>）をグループの全員で一緒に鑑賞するよう指示し

---

<sup>5</sup>Hanna, W., Barbera, J., Jones, C.: *TOM & JERRY: Sci-Fi & Fantasy type episodes*[DVD], Warner Home Video (2003).

た。(2) Sociable Spotlight の実装されたインタラクティブメディアの設置されている会話ブースに移り、各自テーブルの周りに着席するように指示した。このとき座席の位置は図5.1と同様の配置とし、被験者の着席する序列は自由とした。(3) 実験者はグループ全員に対して「先程ご覧頂いたビデオのストーリーを、最初から最後まで、時系列に沿って話して下さい。隣の人と相談し、皆で協力しながら、できるだけ細かく正確に出来事を並べて下さい。このグループがどのくらい細かくストーリーを正確に話していたのか、こちらのビデオカメラで記録しています。」と教示した。(4) 実験者の合図によって5分間の対話を行うように被験者に指示した。(5) テーブルに Sociable Spotlight を投影し、会話開始の合図をした。(6) 5分後、実験者の合図で実験を終了した。なお、被験者には Sociable Spotlight の移動する仕組みを一切知らせていない<sup>6</sup>。

実験環境は次のとおりである。会話中は Sociable Spotlight をテーブル上に投影した。実験中の会話ブースの様子を2台（天井とテーブル横に設置）のビデオカメラで録画した。天井のビデオカメラは被験者から見えにくい位置に設置した。テーブル横のビデオカメラは着座した被験者の目線と同じ高さになり、どの被験者からもよく見える位置にある。実験中は照明を落とし、太陽光をカーテンで遮るなどして、Sociable Spotlight の方に被験者の視線が集中するようにした。

### 5.4.3 分析手法

分析の方法は次のとおりである。実験で収録したインタラクション・ログデータを質的フェーズと量的フェーズ、統合フェーズの3つのアプローチで分析した。具体的には(1) 5.4.4の「質的フェーズ」においてエージェンシーの形成と関わり深い発話行動を同定し、(2) この発話行動の観察されたグループの Sociable Spotlight から受けた影響を5.4.5の「量的フェーズ」で量的な観点から評価し、1の質的フェーズの解釈を妥当化する。そして(3) 1の質的フェーズにおいてエージェンシーが形成されにくいと捉えたグループと比較を行い、着目した発話行動が得られなかつ

<sup>6</sup>実験終了後のデブリーフィングにて明らかにした。

表 5.3: 抽出の手順とエージェンシー付与の判断基準

手順	内容
B-1	Sociable Spotlight の外観や動作を議論する場面を抽出
判別 1	B-1 であれば Sociable Spotlight は話題として会話に取り込まれていると判断
B-2	B-1 が見当たらない場合, Sociable Spotlight にアドレッシングされた発話を抽出
B-3	B-2 の発話の中で隣接対が予期される発話を抽出
判別 2	B-3 に被験者が誰も答えずに会話が進行した場合は「受け手」としてのエージェンシーが付与されたと判断
判別 3	B-1, B-2 が見当たらない場合, 会話の流れの中で利用される価値が見出されなかったと判断

たグループでは Sociable Spotlight から受ける影響が異なることを示す (5.4.6 の「統合フェーズ」)。そして, 1~3 の分析を横断的に理解し, 着目した発話行動が Sociable Spotlight のエージェンシーの構成に関連していたことを示す。

#### 5.4.4 観察結果—質的フェーズ

実験で収録したデータをまず質的に分析した。

##### 抽出手順

表 5.3 の手順に従い, 収録したビデオデータから本観察で取り上げたい (5.4.1 の「観察の概要」で言及した) 3つの発話行動を抽出した。ただし, アドレッシング

グ<sup>7</sup>や隣接対<sup>8</sup>に関する定義を注釈のようにした。

### グループ1

グループ1の収録データから B-2 に該当する発話行動を取り出した。次のような場面である。

ビデオのストーリーを一通り想起し終えた被験者が共同想起対話の終了を隣の被験者に呼びかける。続いて同被験者は Sociable Spotlight を指差しながら、共同想起対話の終了を今度は Sociable Spotlight に対して呼びかける。このような問いかげに Sociable Spotlight は全く反応しなかった（例えば、移動したり、形や色を変えたり、発話したり、など明確に確認できる変化がなかった）。この Sociable Spotlight の振る舞いを確認した被験者らはこのあと、今までの想起方法や想起してきた会話の内容が本当に正しいのかどうか皆で議論し始めた。

つぎに着目している場面のトランスクリプトを示す。

- 01 被験者 A:     で 終わりか↑
- 02 被験者 B:                     終わり 終わりだ
- 03 (3.0)
- 04 被験者 B:     huh
- 05 (2.0)
- 06 被験者 A:     終わりですね これで
- 07 (被験者 A の発話のアドレス先が被験者 B から Sociable Spotlight に移行)
- 08 被験者 A:     初めから 最後まで
- 09 被験者 B:                                     huh
- 10 被験者 C:     んつと:: もうちよつと細かく

<sup>7</sup>ある人物や物体に発話を投げかける・呼びかけるという行為を指すこととする。ただし、Sociable Spotlight に触れる・指差すなどの非言語的行為のみの場合は含まない。

<sup>8</sup>同一話者・受け手の有無に関わらず、先行する発話と後続する発話が対を成しているように見える発話対（質問に対する応答、要請・依頼に対する受諾や拒否など）を指すこととする。

- 11 被験者 A: huhu  
12 被験者 B: fufu  
13 被験者 C: 喋れって事なのかな  
14 被験者 B: え どのぐらい  
15 被験者 B: 細かく喋ればいいのか分かんかったから  
16 被験者 B: 最初めっちゃ細かく言ってやんやけど

分析手順に従い、Sociable Spotlight のエージェンシーについて探った。06 行目～08 行目の「終わりですね これで 初めから 最後まで」という発話は Sociable Spotlight に発話のアドレス先が向けられており、B-2 に該当する。また、この B-2 の発話は別の参加者に対する質問であり、別の参加者からの回答が予期されるため B-3 に該当する。そして、10 行目～13 行目で別の被験者 C が別の質問をしており、06 行目～08 行目の質問に被験者は誰も返答しないまま、会話が進行していることが分かる。従って、誰かがその質問に返答していると考えられる。ここでは、その役割は Sociable Spotlight に与えられていると捉えた。以上のことを考慮し、Sociable Spotlight に「受け手」としてのエージェンシーが形づくられたと判断した。

## グループ 2

グループ 2 の収録データから B-1 に該当する発話行動を取り出した。そのトランスクリプトを以下に示す。

- 01 被験者 A: 割れんかった  
02 被験者 B: そうそうそうそう  
03 (Sociable Spotlight が被験者 A からテーブル中央に移動)  
04 被験者 C: おっ  
05 被験者 B: あっ なんか動いた  
06 被験者 A: なんか 動き悪いん？

- 07 被験者 C: どうなったら来るん これ  
08 被験者 B: よく分かんない

着目している場面は次のような状況であった。同じ場所で約2分間以上停滞していた Sociable Spotlight が突然移動し始めると (03 行目), 一人の被験者が Sociable Spotlight の方に視線を向けて「あっ, なんか動いた」と発話し (05 行目), 今までの共同想起対話が中断されるという現象が起きた。そのあと, 同グループでは Sociable Spotlight が移動した原理やその理由について, 皆で議論し始めた (06~08 行目)。このように「会話の話題」として Sociable Spotlight が多人数会話に取り込まれる場面が見られた。

### グループ3

グループ3の収録データから B-2 に該当する発話行動を取り出した。次のような場面である。

テーブル上で揺れ動く Sociable Spotlight を手で移動させようとする振る舞いを行う被験者が1名観察された。これは実験開始直後に観察されたもので, 同被験者は Sociable Spotlight に手を添えながら「そっちに居て」と発話している。また, このグループの別の被験者2名は先の被験者と Sociable Spotlight とのやり取りを目撃し, そのあと「笑い」を表出している。

つぎに着目している場面のトランスクリプトを示す。

- 01 (Sociable Spotlight が被験者 A から被験者 C に移動)  
02 被験者 A: そっちに居て  
03 被験者 B: [ huh ha  
04 被験者 C: [ huh ha ha  
05 (共同想起対話の開始)

分析手順に従い, Sociable Spotlight のエージェンシーについて探った。02 行目

表 5.4: 各グループのエージェンシー形成時の会話の流れとその形成速度

グループ	会話の流れ	利用のされ方	形成の速度
1	B-2, B-3	受け手	○
2	B-1	話題	△
3	B-2, B-3	受け手	◎
4	無し	無し	△

の被験者 A が Sociable Spotlight に手を添えながら「そっちに居て」と発話した箇所を B-2 とした。また、02 行目の発話は隣接対を予期する発話（要請）であるため B-3 とした。そして 03 行目、04 行目において被験者 B、被験者 C は被験者 A の 02 行目の行為に対して「笑い」を表出し、会話を先に進めている（05 行目以降）。このように B-3 に一人の被験者も返答しないまま会話が進行したことが分かる。従って、誰かがその要請に返答していると考えられる。ここでは、その役割は Sociable Spotlight に与えられていると捉えた。以上のことを考慮し、Sociable Spotlight に「受け手」としてのエージェンシーが形づくられたと判断した。

#### グループ 4

グループ 4 の収録データから B-1 や B-2 が見当たらなかった。とくに、同グループでは被験者は互いの顔だけを注視しており、Sociable Spotlight の方に視線を向けることが困難な状況にあった。このような状況から、同グループの被験者たちがその会話の流れの中で Sociable Spotlight を利用できるタイミングは無かったと判断した。

#### まとめ

各グループのエージェンシー形成時の会話の流れとその形成速度を表 5.4 にまとめた。ここでグループ 3 は実験開始直後に Sociable Spotlight に「受け手」として

表 5.5: コレスポネンス分析で使ったラベル

ラベル	内容
HB1	話し手としての発話 (通常発話)
HB2	聞き手としての発話 (バックチャンネル)
HB3	手の配置を調整する行為 (腕を組む, 髪を触る, テーブルに触れる, など)
SB1	自分の近くに Sociable Spotlight がある (又は自分の方に近づいてきた) 状況
SB2	別の被験者に Sociable Spotlight がある (又は自分から離れていった) 状況
SB3	テーブル中央に Sociable Spotlight がある (又は中央に近付いていった) 状況

のエージェンシーを形成したため、比較的早期にエージェンシーが形づくられている。そのためエージェンシーの形成速度を◎とした。一方、グループ2とグループ4については Sociable Spotlight を「一種の会話参与者」として用いるタイミングがその会話の流れの中では見当たらなかった。従って、エージェンシーが形成されにくい状態にあったと考えられ、形成速度を△とした。

#### 5.4.5 観察結果—量的フェーズ

グループ3では Sociable Spotlight のエージェンシーが早期に形成された。この解釈が妥当であれば、Sociable Spotlight は一人の会話参与者として、同グループの被験者の発話行動に大きな影響を与える。つぎに、その影響を量的な観点から示した。



### 被験者の発話行動に与える影響

ここでは、次のような分析手順に従って分析を進めた。これは「多人数会話に介在するアニメーションの挙動が会話参加者の発話行動にどのような影響を与えるのか」を調べた [121] の実験で使用された分析手順を参考にしたものである。

**A-1** コレスポネンス分析を行い、Sociable Spotlight の変化と被験者の振る舞いの類似性を調べた。これは

**A-1-A** 被験者 A に対して

**A-1-B** 被験者 B に対して

**A-1-C** 被験者 C に対して、行った。

**A-2** カイ二乗検定を用いて、Sociable Spotlight の変化と被験者の振る舞いの相関関係を調べた。

**A-3** Sociable Spotlight の変化と被験者の振る舞いの時間的な因果関係を調べた。

### 分析 A-1：類似性を調べる—コレスポネンス分析

類似性を調べるためにコレスポネンス分析を使用した。コレスポネンス分析とはクロス統計表の中の行と列の各項目の関連性を調べる統計手法である。ここでは行を Sociable Spotlight の変化、列を被験者の振る舞いとしてクロス統計表を作った。

クロス統計表を作成するために、Sociable Spotlight の変化と被験者一人ひとりの振る舞いを表 5.5 に挙げたラベルに従ってカウント<sup>9</sup>した (表 5.6)。ただし、ここでは通常の発話<sup>10</sup>とバックチャネル<sup>11</sup>を分けてカウントしている。このように通常の発話とバックチャネルを分けてカウントすることで、ここでは便宜的に「話

<sup>9</sup>表 5.5 に該当しない被験者の振る舞いはカウントから除いている。

<sup>10</sup>ここでの通常の発話とは、バックチャネル以外の発話を指すこととする。

<sup>11</sup>ここでのバックチャネルとは、話し手の発話を支える相槌や笑い声のことを指すこととする。

表 5.6: Sociable Spotlight の変化に対する被験者ごとの振る舞い

		HB1	HB2	HB3
被験者 A	SB1	7	0	7
	SB2	12	4	19
	SB3	4	1	3
被験者 B	SB1	10	1	7
	SB2	10	8	2
	SB3	6	2	4
被験者 C	SB1	2	5	1
	SB2	2	14	1
	SB3	2	5	3
被験者合計	SB1	19	6	15
	SB2	24	26	22
	SB3	12	8	10

し手としての発話 (HB1)」と「聞き手としての発話 (HB2)」を区別した。こうして得られた表 5.6 の値を元にコレスポンデンス分析を行った。

#### 分析 A-1-A : 類似性を調べる—被験者 A

被験者 A のコレスポンデンス分析結果を散布図で表したものを図 5.4 に示す。このグラフは変数間の類似性を可視化したものであり、関連性の強い項目同士はより近い位置にプロットされる。

被験者 A は HB3 と SB2 が近い位置にある。ここから、被験者 A は Sociable Spotlight が別の被験者にある (あるいは自分から離れていった) 状況 (SB2) と、手の配置を調整するという振る舞い (HB3) が関連していることが分かる。

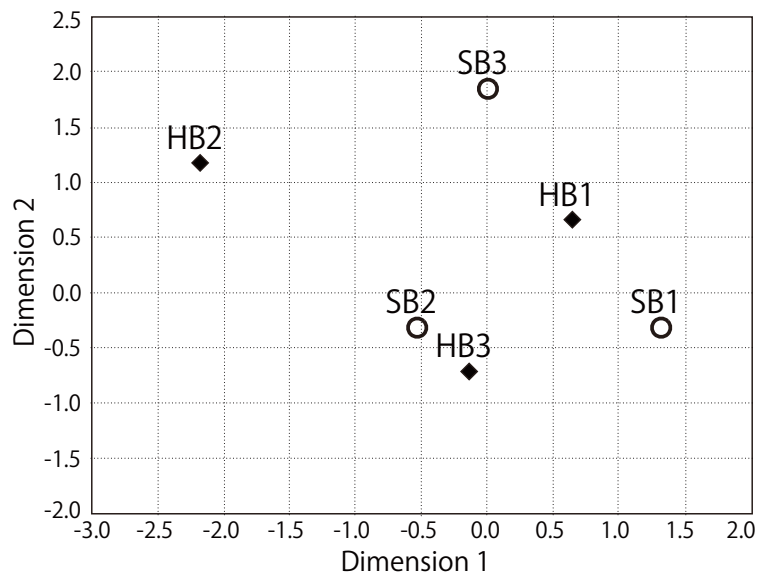


図 5.4: 被験者 A のコレスポンデンス分析結果

#### 分析 A-1-B：類似性を調べる—被験者 B

次に被験者 B について示す (図 5.5)。被験者 B は、HB1 と SB1、HB2 と SB2、HB3 と SB3 が互いに近い位置にある。ここから、自分の近くに Sociable Spotlight がある状況 (SB1) と「話し手としての発話 (HB1)」が関連していること。また、別の被験者に Sociable Spotlight がある状況 (SB2) と「聞き手としての発話 (HB2)」が関連していること。さらに、テーブルの中央に Sociable Spotlight がある状況 (SB3) と手の配置を調整する行為 (HB3) が関連していることが分かる。

#### 分析 A-1-C：類似性を調べる—被験者 C

被験者 C も被験者 B と同様に、HB1 と SB1、HB2 と SB2、HB3 と SB3 が互いに近い位置にある (図 5.6)。

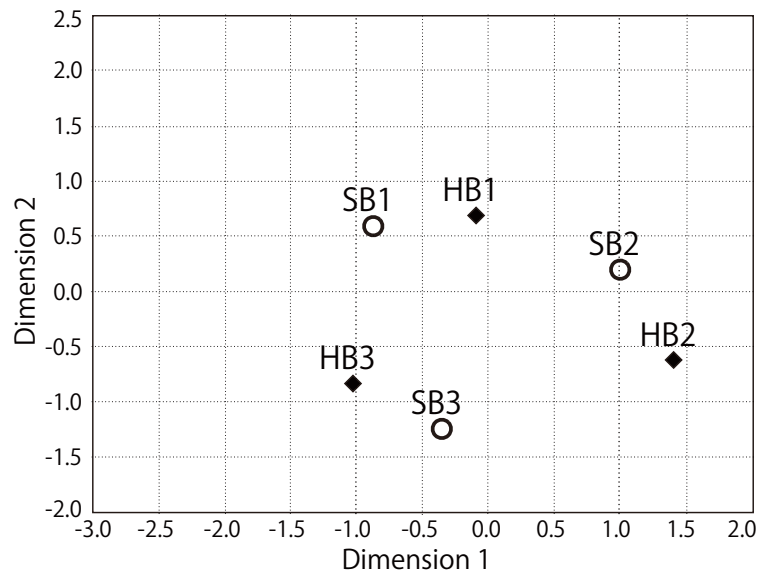


図 5.5: 被験者 B のコレスポンド分析結果

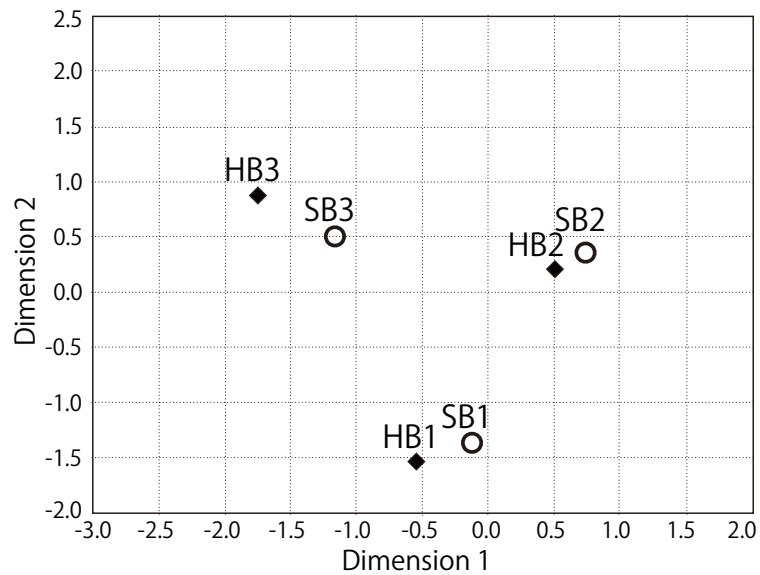


図 5.6: 被験者 C (右) のコレスポンド分析結果

### 被験者 B と被験者 C の振る舞いの類似性

ここで被験者 B と被験者 C の結果を比較すると、Sociable Spotlight の変化に対する被験者 C の振る舞いの対応関係が被験者 B と同一であることが分かる。座標の位置ではなく HB1 と SB1, HB2 と SB2, HB3 と SB3 という組み合わせが等しい。これは非常に稀な出来事である。

表 5.6 に示す通り、Sociable Spotlight の変化と被験者の振る舞いの対応関係は全部で 9 通り (HB1 と SB1, HB2 と SB1, HB3 と SB1, HB1 と SB2, HB2 と SB2, HB3 と SB2, HB1 と SB3, HB2 と SB3, HB3 と SB3) 存在する。加えて、Sociable Spotlight と無関係に行為を調整するという組み合わせがある。このように組み合わせが多様に存在するにも関わらず、Sociable Spotlight の変化に対して自分の行為をどのように関連付けるのかという「やり方」が 2 人の被験者で共通している。

このような結果から、Sociable Spotlight の変化に合わせて被験者らの行為をそのようにしむける選択圧 (selection pressure[93]) のようなものがこのグループ内において働いていると考えられる。その可能性を調べるために、カイ二乗検定を用いて Sociable Spotlight の変化と同グループの被験者の振る舞いの相関関係を調べた。

### 分析 A-2：相関関係を調べる—カイ二乗検定

Sociable Spotlight が自分の近くにある状況と別の被験者にある状況の全員の被験者の「話し手としての発話」と「聞き手としての発話」の頻出回数をカイ二乗検定した。

表 5.6 に示す通り、このグループの被験者全体としては Sociable Spotlight が自分の近くにある状況 (SB1) では「話し手としての発話 (HB1)」を 19 回繰り出し「聞き手としての発話 (HB2)」を 6 回発している。一方、Sociable Spotlight が別の被験者にある状況 (SB2) では「話し手としての発話 (HB1)」を 24 回繰り出し「聞き手としての発話 (HB2)」を 26 回発している。

この結果をカイ二乗検定したところ、Sociable Spotlight が自分の近くにある状況

と別の被験者にある状況では「話し手としての発話」と「聞き手としての発話」の頻出する割合に有意な差があることが認められた ( $\chi^2=5.342$ ,  $p=0.0208$ ,  $d.f.=1$ )。

このような結果から、主に SB1 と SB2 という Sociable Spotlight の変化が、HB1 や HB2 という被験者の発話行動を促していると考えられる。しかし、Sociable Spotlight の変化が被験者の振る舞いに制約を与えたのか、被験者の振る舞いが Sociable Spotlight の変化に影響を与えたのか（どちらが原因でどちらが結果なのか）が十分に明らかでない。そこで、被験者の振る舞いを時系列に沿って分析し、両者の因果関係を調べた。

### 分析 A-3：時間的因果関係を調べる

Sociable Spotlight の変化が被験者の発話行動に制約を与えていることを明らかにするため、さらに時間的な因果関係を調べた。音響的な影響<sup>12</sup>が混入することを最小限に抑えるために、沈黙<sup>13</sup>における次の発話者がどのようにして選択されるのかを調べた。

ELAN<sup>14</sup>を用いて、各被験者の発話区間と沈黙区間、Sociable Spotlight の位置をアノテーションした。その結果を各沈黙区間ごとに整理したものを図 5.7 に示す。ここでは、2名の判定者によるアノテーション作業を実施し、判定者間の評価の一致率を計算した。共同想起対話の開始から終了までの約5分間を判定区間とし、Cohen のカッパ係数を算出した。ここでは、判定区間を5秒毎に分割し、合計61個に分割されたそれぞれの区間毎において2名の判定者間の評価を比較している。また、判定対象を被験者 A、被験者 B、被験者 C の発話区間および沈黙区間とした。カッパ係数は0.664で良好な一致率であった。よって、図 5.7 のアノテーション結果の信頼性は確保された。Sociable Spotlight の位置や変化についてはビデオデータから定量的に判断した。

<sup>12</sup>名前による呼びかけや指名などにより特定の参加者が話し手になることを強制する機能。

<sup>13</sup>現在の話者が次の話し手を指定しなかったことで誰かが順番をとって発話するまでに生じる沈黙に着目している。

<sup>14</sup>細馬宏通：会話分析・ジェスチャー分析研究者のための ELAN 即席入門，<http://www.12kai.com/elan/index.html> (2009)。

区間	被験者の発話状況	次話者の選択	SSL交替規則
S1	<p>A_utterance B_utterance C_utterance Spotlight (A, C) silence 00分07秒200</p>	<p>聞き手 B A 発話者 C 聞き手 00:07.200</p>	該当
S2	<p>A_utterance B_utterance C_utterance Spotlight (C) silence 00分14秒600</p>	<p>聞き手 B A 聞き手 C 発話者 聞き手 00:14.600</p>	該当
S3	<p>A_utterance B_utterance C_utterance Spotlight (B) silence 01分47秒400</p>	<p>発話者 B A 聞き手 C 聞き手 聞き手 01:47.400</p>	該当
S4	<p>A_utterance B_utterance C_utterance Spotlight (A, B) silence 03分55秒700</p>	<p>発話者 B A 聞き手 C 聞き手 聞き手 03:55.700</p>	例外
S5	<p>A_utterance B_utterance C_utterance Spotlight (A) silence 04分43秒700</p>	<p>聞き手 B A 発話者 C 聞き手 聞き手 04:43.700</p>	該当
S6	<p>A_utterance B_utterance C_utterance Spotlight (B) silence 05分07秒000</p>	<p>発話者 B A 聞き手 C 聞き手 聞き手 05:07.000</p>	該当

図 5.7: Sociable Spotlight の動作と沈黙区間における話者選択

ここで、図5.7に示した各沈黙区間 (S1-S6) のセル内上部 A\_utterance~C\_utterance はそれぞれ被験者 A~被験者 C が発話をしてきた区間<sup>15</sup>を示している。その下部の spotlight は Sociable Spotlight の停滞区間を示している。例えば A という注釈が付けられている区間は被験者 A から約 40cm 以内の領域に Sociable Spotlight が位置していたことを示している。最下部の層 silence はグループの沈黙区間<sup>16</sup>を示している。

ここで、各沈黙区間における Sociable Spotlight の変化と被験者間の時間的因果関係に着目する。会話中の沈黙区間は全部で6箇所あり、その内の5箇所ですべて「Sociable Spotlight を媒介にした発話交替規則 (SSL 交替規則と呼ぶ)」がみられた (図5.7 右)。この規則とは、沈黙区間に Sociable Spotlight がより長く自分の付近に停滞している (もしくは自分から離れていった) 状況にある被験者が、沈黙を破って次の発話者となるという傾向である。なお、このような振る舞いは特定の被験者だけではなく同グループの全ての被験者から引き出されている。

図5.7において、最初の沈黙区間 S1 では沈黙区間の直前まで被験者 A に Sociable Spotlight が停滞していた。3 秒程の沈黙の後、被験者 A から短い発話を得られている。と同時に他の被験者の発話と重複することがない。また2番目の沈黙区間 S2 では、被験者 C に持続的に Sociable Spotlight が停滞していた。3 秒ほどの沈黙の後、被験者 C の発話を得られた。と同時に他の被験者の発話と重複することがない。さらに3番目の沈黙では、その沈黙区間において被験者 B に Sociable Spotlight が近寄り、そのあと被験者 B から発話を得られている。と同時に他の被験者の発話と重複することがない。以降5番目と6番目の沈黙区間においても、同様の傾向がみられた。

このように、5.4.4 の「質的フェーズ」において、Sociable Spotlight のエージェンシーが早期に形成されたと解釈したグループ3に着目すると、Sociable Spotlight は同グループの被験者たちの参与役割を決定づける要因の一つ<sup>17</sup>として、被験者間

<sup>15</sup> 笑い声やバックチャネルなどを除く発話区間。

<sup>16</sup> 3 秒以上の無音 (笑い声やバックチャネルなども含め一切の発話のない) 区間。

<sup>17</sup> 4 番目の沈黙はこの傾向に従わないため、他の要因によっても沈黙における次の話者は決定されている。



の発話連鎖に影響を与えていることが分かった。これは5.4.5の「量的フェーズ」の冒頭で言及した予想と一致しており、質的フェーズでの解釈を妥当化する一つのエビデンスといえる。

ここで興味深い点は、早期にエージェンシーの形成されたグループでは、Sociable Spotlightが参与者たちの発話連鎖を制約し、方向づけるようなアーティファクトとして機能しているという点である。そして、このSociable Spotlightが一種の会話参与者として発話連鎖に影響を与え得るということは、同グループとは対照的なエージェンシーの形成されにくいグループでは、Sociable Spotlightの動作に対して異なった発話連鎖が生ずると考えられる。次節では、エージェンシーの形成された/されにくいグループ間に生じる発話連鎖の差異を調べ、質的フェーズ/量的フェーズでの解釈をサポートするエビデンスを探った。

#### 5.4.6 観察結果—統合フェーズ

最後に、エージェンシーが形成されたグループとされにくい状態にあったグループの被験者の発話行動を比較し、Sociable Spotlightとのインタラクションの全体像を把握した。ここではPC内に記録されたログデータから、被験者の音声データとSociable Spotlightの移動履歴を利用した。

一人ひとりの被験者の口元に取り付けたマイクロフォンのインプットレベルが閾値<sup>18</sup>を上回った瞬間を当該被験者の発話開始箇所とし、下回った瞬間を発話停止箇所とした。また、Sociable Spotlightのテーブル上の座標と各被験者の座席との間を結ぶ直線距離が16ミリ秒前と比べて短くなった場合、Sociable Spotlightが当該被験者に「近づいている」とし、長くなった場合は「離れている」とした。このログデータからSociable Spotlightの動作に対して被験者のとった発話行動を表5.7に示す4つのパターンA~Dに分類し、それぞれの生起する頻度をグループ毎にカウントした。グループ1~4の結果をグラフにまとめたものを図5.8に示す。

ここで、「受け手」としてのエージェンシーが形成された2つのグループ（グルー

<sup>18</sup>ここでは40dBとした。

表 5.7: 被験者の発話行動の分類

パターン	内容
A	Sociable Spotlight が被験者に近づく場合に同被験者が発話を停止する
B	Sociable Spotlight が被験者に近づく場合に同被験者が発話を開始する
C	Sociable Spotlight が被験者から離れる場合に同被験者が発話を停止する
D	Sociable Spotlight が被験者から離れる場合に同被験者が発話を開始する

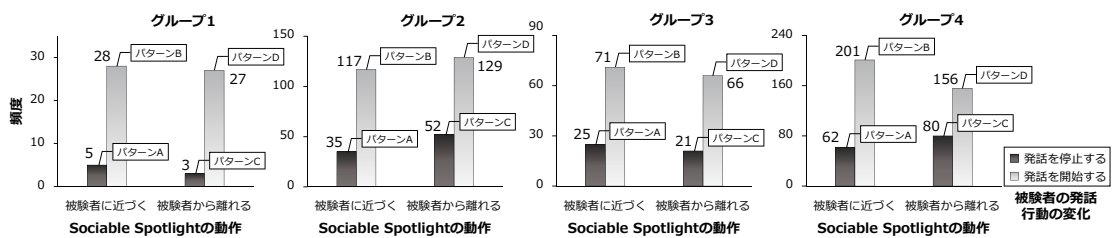


図 5.8: Sociable Spotlight の動作に対する被験者の発話行動の変化 (グループごとに算出)

グループ1とグループ3)と、エージェンシーの形成されにくい状態にあったグループ(グループ4)を比較する。前者ではパターンCよりもパターンAの方が頻度が多いのに対し、後者ではパターンAよりもパターンCの方が頻度が多い。このように、エージェンシーが形成されたグループとされにくい状態にあったグループでは、パターンAとパターンCの生起する比率が反転している。ここから「受け手」としてのエージェンシーが形づくられるグループとエージェンシーの形成されにくい状態にあったグループとの間で相反する傾向が得られた。これは5.4.5の最後の段落で言及した予想と一致しており、質的フェーズ/量的フェーズにおける議論をサポートするエビデンスの一つといえる。

このような相反する結果が得られた理由として、次のことが考えられる。5.4.5の「カイ二乗検定」で示したように、エージェンシーが比較的早期に形成されたと解釈したグループ3では、Sociable Spotlightの変化に合わせて「話し手としての発話（通常の発話）」が引き出されるとともに、バックチャネルや笑い声などの「聞き手としての発話」も一緒に引き出されている。このような出来事は一人ひとりの発話のタイミングを他の参与者と重複させ、重複発話の生起頻度を増加させる。また、この重複発話の生起はパターンAの頻出回数を増加させる。

パターンAの「Sociable Spotlightが被験者に近づく」という動作は、同被験者の発話継続時間がグループ内で最も短くなったことを示している。また「発話を停止する」という条件は、それまで同被験者が発話中であった場合にカウントされる。このように、発話中にもかかわらずSociable Spotlightが近づいているため、5.4.2の「移動関数の調整」の中で例示したように、他のメンバは同被験者の発話中に発話をしている。このような理由から重複発話が多い場合にパターンAは生起しやすい。

一方、パターンCはパターンAとは対照的に、一人ひとりの発話のタイミングが別の会話参与者と交互に生じる場合に生起しやすい。一般的な会話では、一人ひとりの発話は交替に生じることが知られている [73]。その意味で、Sociable Spotlightからの影響を受けにくい場合では、一般的な会話と同様に、パターンAよりもパターンCの方が生起しやすい。

このような理由から、エージェンシーが形成されやすい場合にパターンAの生起頻度が多くなり、それ以外ではパターンCの生起頻度が多くなったと考えられる。

## 5.5 全体総括と今後の課題・展望

本研究では、Human-Agent Interaction研究が前提としてきた「モノのエージェンシーはそのモノ自体に備わる」という考え方から脱却し、人と人の社会的な関わりの中で生み出され、人と人の社会的な関わりを方向づけるアーティファクトとして、モノのエージェンシーを議論することを目的とした。見かけや形状とい

う図形自体の特徴からではなく、人とモノ（図形）との因果的および空間的な関わりの中で、モノのエージェンシーは構成されるという観点が本研究の主張である。そのため、社会的な関わり（人からの語りかけや人による解釈）によって形づくられるエージェンシーに着眼する実験システム（Sociable Spotlight）を提案し、2種類の社会的なインタラクションの場に介在させる2つの観察を通じて、Sociable Spotlightのエージェンシーが社会的な関わりの中で構成されていることを論じた。

1つ目の観察では、複数名の来場者が互いに対話するという状況がSociable Spotlightのエージェンシーを構成するうえで無視することのできない要素であることを述べた。この結果をうけて、2つ目の観察では、被験者間の発話のやり取りの細部に着目してSociable Spotlightのエージェンシーがどのように構成されていくのかを示した。

ここで明らかになったのは、会話の流れの中で、一種の会話参与者として（発話の受け手として）Sociable Spotlightの動作を会話参与者が利用する場合があります、また、そのようにして社会的に構成されたSociable Spotlightの「受け手としてのエージェンシー」が人と人のインタラクションの様相を変容させている点である。具体的には、まず5.4.4の「質的フェーズ」において、隣接対と呼ばれる発話と発話のやり取りの中で、Sociable Spotlightの「受け手としてのエージェンシー」が形づくられることに着目した。表5.4では、そのようなやり取りの中でエージェンシーが形成された/形成されないグループの二つを区別した。そして、前者のグループでは、Sociable Spotlightの動作が被験者の発話連鎖に影響を与えていることを示した（5.4.5の「量的フェーズ」）。また、エージェンシーの形成されたグループと形成されないグループでは、Sociable Spotlightの動作による被験者の発話連鎖に与える影響が異なることを明らかにした（5.4.6の「統合フェーズ」）。

このように相反する結果が得られる理由は二つ考えられる。一つは、Sociable Spotlightの設計にミニマルデザインを用いた点である。ミニマルデザインにより、被験者間でSociable Spotlightの動作に対する解釈が異なる。その解釈の違いが結果に影響を及ぼしたと考えられる。もう一つは、Sociable Spotlightに関わる際の被験者の心理的な要因がある。しかし今回の実験では、観察の着眼点を被験者の

発話行動に絞った。被験者の心理的な要因の分析は、本論文の範囲では問題としない。被験者の統制や実験課題の選定、より詳細な分析や巧妙な心理実験については他の機会に譲ることとする。

もう一つは、Sociable Spotlight が次世代インタフェースとしてユニークな性質を備えていた。実社会で自由に行動している人々から共同行為を引き出すためには技量を必要とする。Sociable Spotlight は人々が自由に行き来する公共施設において、通りがかった複数の来場者を一つの共同行為に引き留めていた。また、多人数会話の場面では、人と人の発話タイミングを適度にバランスして発話と発話の偏りを避けたり、沈黙において特定の参加者のみから発話を引き出していた。このように、ミニマルデザインに基づくシンプルなアニメーションであっても、共同行為を引き出すインタフェースとして働く。ここでは、とくにその目的や使用状況を明確に定めていないため、参加者の自由な解釈や関与の仕方によって、様々なシーン・目的で機能していた。このように人々の社交空間のあり方を変容させるような次世代インタフェースとしての可能性についても調べたい。



## 第6章 説得性の社会的構成に着目したメディア技術のHRIへの応用例

### 6.1 はじめに

私たちは大地の上を歩くと同時に、大地が私たちを歩かせている。一度、そうした視点から私たちの身の回りのことを見渡してみるのも面白い。街の中をなにげなく歩く時に、「街を歩いているのは、この私だ」と自分を中心に考えてしまいがちだけれど、その歩道や街の看板、駅に向かう人の流れは、私たちの歩くという行為を制約し方向づける。その意味で、なにげない一歩とそれを支える大地や街並みとが一緒になって「歩行」という行為を生みだしているといえる。

これは「歩く」という行為に限られない。なにげない落書きも、なにげない手の動きとともに、その鉛筆と紙との間にある摩擦に支えられながら生まれるものだろう。なにげなく生まれたラインは、次の描くという行為を制約し方向づけている。それと、人前で何かを話そうとするとき、その相手が無表情なときには、とても話し難いということがある。なにげなく言葉を繰り出そうとするとき、その相手は私たちの発話が向かう対象であると同時に、私たちの発話の内容を制約し方向づける。つまり、その聞き手のうなずきや視線の動きに支えられながら、一緒になって発話を生みだしているといえる。

一方で、これまでの発話生成システムや音声合成システムの場合はどうだろう。スイッチをオンにすると、そのスピーカーからは確かに合成音は聞こえてくる。しかし、その合成音は「宛名」を伴ったものではない。また、その発話は聞き手の存

在を予定したもので、その状態に配慮したものでもない。その意味で、私たちがテキストの読み上げソフトを使用する際には、そうした「宛名のない合成音」の聴取を一方的に強いられてきたともいえるだろう。このような状況にあつて、ソーシャルなロボットのフレームワークを援用することで、これまでの発話生成システムのあり方を再構築できないものかと考えた。

ソーシャルなロボットは、その視線や姿勢の調整機構を備えており、比較的容易に「誰にアドレスを向けるか」という宛名性を備えさせることができる [37, 46]。また、そのセンサやカメラなどによって聞き手の状態を発話生成に反映させることも可能だろう。会話エージェントの分野では、ユーザの視線方向を検出して、聞き手の興味や対象に合わせて話題を変更するというエージェントが開発されている [120]。加えて、ロボットの適応学習機構を利用して、その発話方略を聞き手の個性に適応させることも考えられる。

もう一つの期待は、その発話や合成音に対する私たちの構え (stance) に影響を与えることができるのではないかということである。テキストを機械的に読み上げるだけのシステムでは、そこに主体性や志向性を感じることはないように思われる。一方で、聞き手に対する宛名性や相互行為的な調整を含む発話は、私たちの志向的な構えを引き出す可能性がある。このことで、発話の「説得性」を増すことができれば、この発話生成システムの応用範囲をさらに押し広げることになるだろう。

本研究では、これらの可能性を探る一環として、本研究のプラットフォームとなるプロトタイプが発話生成システム Talking-Ally を設計し、その実装を行った。また、このプロトタイプが発話生成システムを使用して、被験者との相互行為的な調整に基づく発話生成システムの基本動作を微視的に調べるとともに、そこで生成された発話に対する聞き手側からの印象評定を行った。さらに、被験者のインタラクションログを用いた行動分析から、生成された発話の説得性の効果を確かめた。本章では、これらの概要と実験結果、考察について述べる。



## 6.2 研究の背景

### 6.2.1 発話の非流暢性に対する解釈

私たちはなぜ言い直すのか、なぜ言い淀むのか。自然な発話 (spontaneous speech) における発話の非流暢性 (disfluencies) は、その要因を話し手の内部のメカニズムに一方的に帰属させて、説明されることが多い。例えば、Levelt らは、発話者の発話におけるプランと行為との乖離 (slips)、そしてその修復プロセスとして、言い誤りや言い直しなどの現象を説明する [49]。こうした状況は、「サイモンの蟻」と呼ばれる議論 [77] に当てはめて考えることができる。

「砂浜を歩く蟻の残した足跡は、なぜ複雑な絵模様を描くのか」ということを考えるとき、私たちはその要因を蟻という個体に一方的に帰属させて考えやすい。その一方で、その複雑さを生み出す要因をその蟻を取り囲んでいる砂浜の起伏に帰属させる考え方もある。もうすこし妥当な説明としては、その間にあるものだろう。つまり、その複雑な絵模様は蟻の内部の機構とそれを取り囲んでいる砂浜の起伏との関わりの中で、結果として生じたものという見方である。

発話の非流暢性に関しても、同様の見方ができるのではないか。話し手の側に一方的にその要因を帰属させるにも無理がある。一方で、聞き手側だけに、その要因を求めることもできない。むしろ、話し手と聞き手との相互行為的な調整の中で、結果的に生み出されたもの、組織されたものと捉えても面白い。そうすることで、発話の非流暢性に対する見方に広がりを持たせることができる。

### 6.2.2 会話における「宛名性」と「聞き手性」

話し手に生じる発話の非流暢性に対して、話し手と聞き手の相互の調整の中で結果的に生み出されたものであると捉えることは発話の非流暢性の現象に留まらず、発話そのものの意味に対してもいえる。

こちらに近づいてくる知り合いに「おはよう」と挨拶をするとき、その相手が何事もなく過ぎ去ってしまったらどうだろう。その「おはよう」の意味は宙を彷徨

徨ってしまう。その意味で、私たちの発話は相手からの応答的な発話があつて、はじめて明確な意味が立ち現われてくるように思える。一方的に発話を繰り出しただけでは、その発話の意味は完結しない。

また言葉というのは、本来はだれかに向けられたもので、かならず宛名を伴ったものだという。これは対話の哲学者として知られる、Mikhail Bakhtinの「宛名性(addressivity)」に関する指摘[52]である。

そして、私たちがいま話し手であるのは、相手が聞き手になっていてくれるからに他ならない。会話の中での話し手や聞き手、傍観者などの参与役割もまた、その会話参与者間での相互行為的な調整の結果として組織される。こうした相互行為的な調整においては、Goodwinの指摘[29]にあるように、聞き手の方もまた「いま、あなたの話をちゃんと聞いてますよ」という「聞き手性(hearership)」の表示が欠かせない。私たちは、その「聞き手性」をリソースとして、発話を組織化しているといえる。

こうした観点で考えると、これまでの発話生成システムは「宛名性」を欠いており、同時にその発話を生み出す際に、聞き手の参加する余地を欠いている。「テキストの読み上げシステム」のスピーカーからは確かに合成音声は聞こえてくるものの、そこに宛名を伴っていない。そういう宛名のない音声の聴取を私たちに強いてきたように思われる。

本研究では、聞き手の存在を予定しつつ、その聞き手との相互行為的調整に基づいて発話の組織化を行う発話生成システムによってどのような印象を聞き手に与えることができるのか。また、従来の発話生成システムとどのような印象の違いが生まれるのかを調査する。発話における非流暢性の生成メカニズムを構成論的に探るためのプラットフォームとなると同時に、聞き手の状態に配慮した、聞き手に優しい発話生成システムを提供することにつながると考えられる。

### 6.2.3 設計的な構えと志向的な構え

その発話生成システムはテキストを機械的に読んでいるだけなのか、それとも聞き手を志向しつつ、聞き手の状態に合わせて、情報を伝えようとしているのか。この「どのように発話が生成されているのか」という発話の生成様式は、その発話に対する聞き手の構え (stance) に影響を与えるものと思われる。

Dennett は、私たちが何か動くものを目にしたとき、その対象に対して、物理的な構え (physical stance)、設計的な構え (design stance)、志向的な構え (intentional stance) のいずれかの構えを取ることを指摘している [11, 12]。これは発話そのものや発話生成システムにも当てはまるものだろう。

例えば、平坦なスピードで、テキストを機械的に朗読した発話に対して、私たちの多くは設計的な構えを取るのではないだろうか。一方、聞き手の状態に配慮して、丁寧になされる発話に対しては、多くの人は志向的な構えを取るように思われる。これらをまず確認することが本研究の目的となる。

同様に、この発話生成の様式に対して、Reeves & Nass の Media Equation に関する議論 [69, 70] を当てはめることができる。音声や応答などのわずかな手掛かり (minimal cues) によって、私たちはコンピュータに対して、礼儀正しさを感じたり、何らかの人格を付与するなど、あたかも人に対するかのような反応 (=対人的な反応) をしてしまう傾向があるという。

本研究で進めている聞き手の状態に配慮する発話生成システムに対してはどうか。その発話に対して、誠実さや優しさ、子どもらしさなどを感じることはあるのか。これらの要素に対して、発話の生成様式がどのような影響を与えるのかは興味深い。

また、こうした議論に加えて、発話生成システムにとってもう一つ重要なことは、その発話は他者を揺り動かすような説得力を備えるものなのかどうか、という点である。Nass らの Media Equation の議論を拡張する形で、Fogg らの説得性を備えたテクノロジー (Persuasive Technology) に関する議論が盛んに行われている [20, 21, 22, 35]。

どのような発話の生成様式が聞き手を揺り動かすような説得力を伴うのか。こ

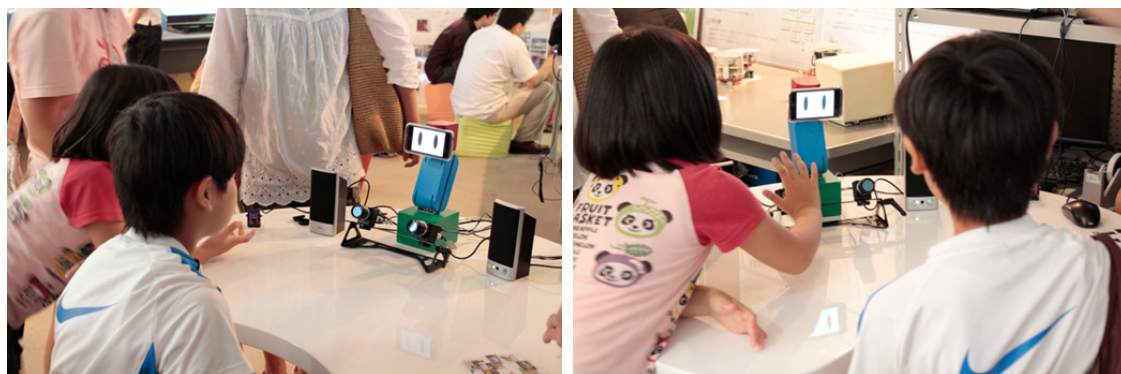


図 6.1: Talking-Ally とのインタラクション

うした側面を明らかにすることも、本研究の狙いの一つである。相手を揺り動かすような説得力の追求に向けて、ロボットが何らかの意図を表出しようとしていることをまず相手に汲み取ってもらえなければならない。ここではその第一ステップとして、聞き手（ユーザ）の志向的な構えを引き出しやすい発話生成システムとはどのような様式かを調べた。

## 6.3 プロトタイプの設計と実装

### 6.3.1 Talking-Ally の構築

6.2節での議論を考慮に入れた発話生成システムを実装するためのプラットフォームとして、図6.1に示すTalking-Allyを構築した。Talking-Allyは3つのサーボモータが内蔵された小型のロボットで、頷く、体を起こす、顔を向けるなどの細やかな動作が可能なロボットである（図6.3）。ここでは、Webから取得したニュース記事（News Text）を読み上げるという発話生成システムをTalking-Allyに実装した（図6.2）。

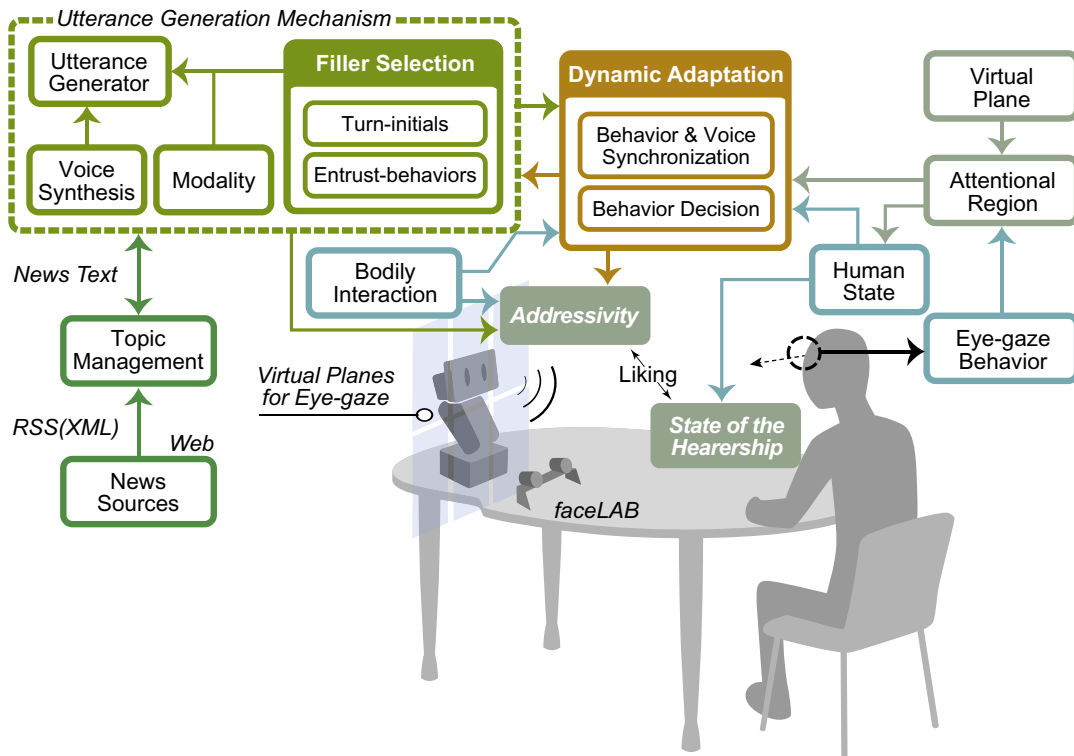


図 6.2: Web 上のニュース記事を読み上げる Talking-Ally のシステム構成

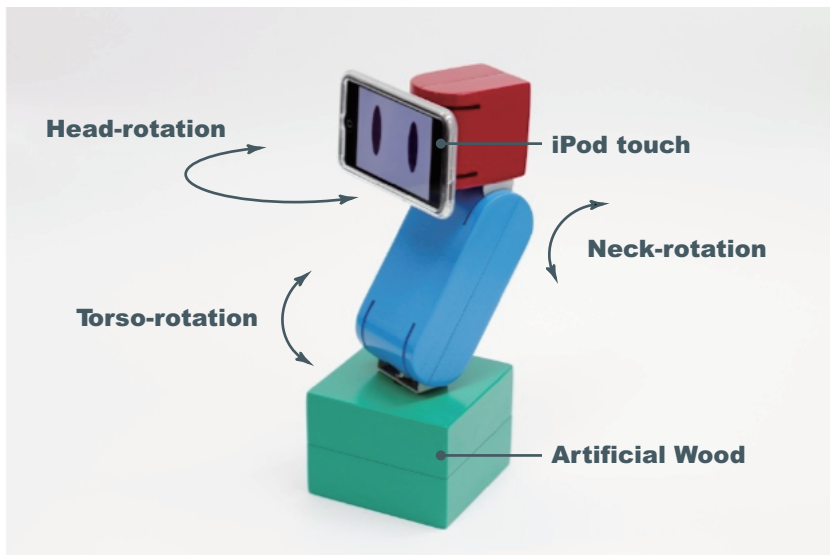


図 6.3: 3つのサーボモータと iPod touch で構成される Talking-Ally の身体

### 6.3.2 発話のデザイン

私たちの発話は誰かに向けられたもので、その宛先となる聞き手の聞き手性をリソースとしながら、少しずつ発話を繰り出している。発話のある構成単位に区切り（フレーズング）、相手の様子（Human State）を伺いながら少しずつ発話しているのだろう。この構成単位に区切られた発話によって、次の発話を行なうべきかどうか判断される。

さらに私たちは、聞き手の関心、うなずき、視線（Eye-gaze Behavior）などの状態をリアルタイムに反映させ、「えーと」「あの一」「それでね」などのターン開始要素（turn-initials）を利用する。ターン開始要素は話し手が聞き手に「今から私が発話しますよ」ということを伝える効果がある。聞き手はこれを受けると、話し手に「あなたの話を聞きますよ」ということを視線などで伝える。また「今日ね、学校でね、先生がね…」などの各構成単位のモダリティ（modality）や、「…なんだって」「…らしいよ」などの発話末のモダリティを利用する。モダリティには、発話内容を話し手の内部で完結させるのではなく、一部を聞き手に委ねることで、聞き手との相互行為を強く意識し、聞き手の参加を促すねらいがある。私たちの会話では、ターン開始要素やモダリティだけでなく、言い淀みや言い直しも多く見られる。

これらを考慮して発話を動的に選択し、会話連鎖を組織化する（図 6.2 Filler Selection）。音声合成には、音声合成エンジン Wizard Voice（ATR-P 社）を利用した。

### 6.3.3 聞き手の状態のセンシング

私たちは聞き手の状態（State of the Hearership）を常に確認し、その「聞き手性」をリソースとして、動的に発話方略を調整している。「聞き手性」の一つである「視線」は、聞き手の志向性を示し「いま、あなたの話を聞いていますよ」という表示と同時に、自分がいま聞き手であるということを話し手に伝える役割がある。このような聞き手の視線に関する認識機構として、頭部・視線システム faceLAB5.0

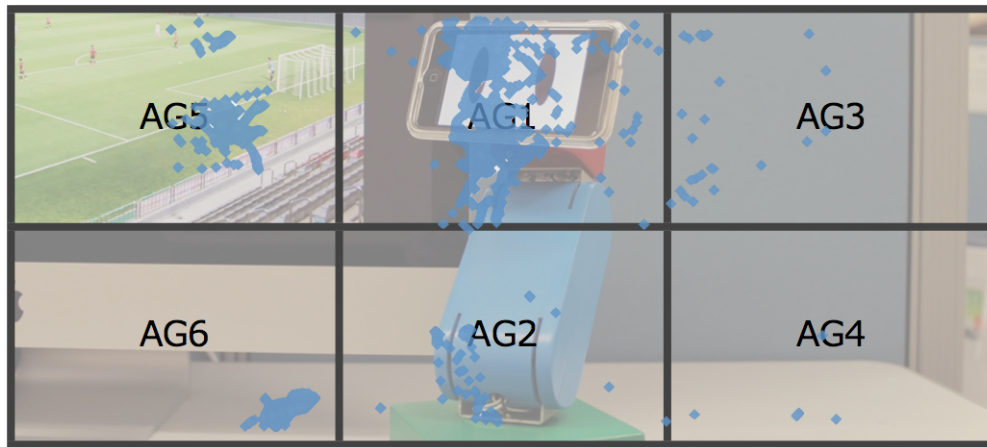


図 6.4: 仮想平面とユーザの視線の推移

(SeenigMachines 社) を用いた。

faceLAB5.0 ではソフトウェア上に図 6.4 に示す仮想平面 (Virtual Plane) を作成し、いま聞き手がどの仮想平面を見ているかを認識できる。図 6.4 の左側に示される仮想平面 AG5, 仮想平面 AG6 が TV への視線に、右側の仮想平面 AG3, 仮想平面 AG4 が周囲への視線に、中央の仮想平面 AG1, 仮想平面 AG2 が Talking-Ally への視線にそれぞれ対応している。AG1~AG6 のどの仮想平面を見ているかという注目度合いの偏りによって、ユーザの関心の対象を推定し、ロボットの振る舞い (Behavior) を調整し、前節の発話方略を動的に選択し組織化する (図 6.2 Dynamic Adaptation)。

表 6.1 に、具体的にどのような要素をニュース文章に付与して発話を生成するのかをまとめた。表 6.1 に示すように、聞き手の注意や興味対象、聞く姿勢によって、ニュース文章に付与される要素が異なる。フレーズされた (ニュース文章中の) 一つ一つの構成単位の前後で、表 6.1 の要素のいずれか一つが構成単位毎に付与されていく。ここでは、turn-initials (TI) や modality (MD), 要求的な発話 (EB) にいくつかのバリエーションを持たせて、その中からランダムに選択されるような実装にしている。表 6.1 はユーザの個性に合わせて最適なバリエーションを選択する前の仮のものであり、この中からユーザの個性に合った最適なバリエーションを見つけ出す機構やその実装については今後の課題とする。

表 6.1: インタラクションの生成方法

聞き手の振る舞い (継続的に見る仮想平面)	ロボットの振る舞い (ニュース文章に付加)
仮想平面 AG1, AG2 (Talking-Ally への注視)	TI1: “あーあ”
	TI2: “あの一”
	TI3: “あのね”
	TI4: “あのさ”
	TI5: “えーっと”
	TI6: “えっとね”
	TI7: “えっと”
	TI8: “えっと一”
	TI9: “ねーねー”
	TI10: “んっと”
	TI11: “んっとね”
仮想平面 AG5, AG6 (別の興味対象への注視)	MD1: “…ね”
	MD2: “…なんだって”
	MD3: “…らしいよ”
	EB1: “聞いてね”
	EB2: “聞いてよ一”
	EB3: “聞いてよね”
	EB4: “こっち見てね”
	EB5: “こっち見てよ”
	EB6: “こっち見てよね”
EB7: “もう一回言うよ”	
EB8: (言い直す)	
EB9: (言い淀む)	



表 6.2: トランスクリプト中の注釈の中で使用したラベルとその内容

ラベル	内容
H	ユーザ
R	ロボット (Talking-Ally)
—forward	姿勢を前へ
—back	姿勢を後ろへ
—nod	うなずき
—human 's gaze	人が向いている方へ視線を向ける
—gaze front	視線を正面に戻す
gaze in	Talking-Ally に視線を向けている
gaze out	Talking-Ally に視線を向けていない

## 6.4 動作例

ここで構築した発話生成システムを用いてユーザとインタラクションした際の動作例を次に示す。以下のトランスクリプト中の注釈の中で使用されているラベルの内容は表 6.2 の通りである。

なお、ここでの動作例においてはユーザ側からの発話は想定していない。被験者の視線の変化を手掛かりにロボットの発話をリアルタイムに変化させる動作をもって、ここでは「相互行為的な調整を含む発話」としている。ユーザ側の発話や身体の動き・ジェスチャなどを考慮に入れた相互行為的な調整は今後の課題とする。

### 6.4.1 インタラクション例 1

インタラクション例 1 では Talking-Ally が言い直しを行った場面を示している。まず、ユーザの視線が得られていない状態で、Talking-Ally が姿勢を前に倒し、話し始めるが、依然としてユーザの視線が Talking-Ally の方に向かなかつたため、再

度同じ発話を行い、ユーザからの視線を得ようとしている。

以下に着目している場面のトランスクリプトを示す。

- 01 ((H: gaze out))  
02 ((R: nod))  
03 (0.8)  
04 ((R: forward))  
05 R: 沖縄警察のうるま署はね  
06 ((H: gaze in))  
07 ((H: gaze out))  
08 ((R: back))  
09 (1.0)  
10 ((R: nod))  
⇒ 11 R: 沖縄警察のうるま署はね  
12 ((H: gaze in))  
13 ((H: gaze out))  
14 ((R: back))

### 6.4.2 インタラクション例2

インタラクション例2では Talking-Ally が発話を躊躇した後で、言い淀みと、言い直しの両方を行った場面を示している。まず、ユーザの視線が得られていない状態で、Talking-Ally は姿勢を前に倒し、話し始めるが、依然としてユーザの視線が得られなかったため、それまでの発話を一時中断し、うなずきと共に「えーっとね」と発話し、ユーザの視線を得ようとしている。さらにユーザは、Talking-Ally 以外の場所を多く見ていたため、「こっち見てよね」という直接的な要求をすることで、Talking-Ally はユーザの視線を獲得しようとしている。その後、先ほど躊躇して言い淀んだ発話を再開している。

以下に着目している場面のトランスクリプトを示す。

31 ((H: gaze out))  
32 ((R: forward))  
33 (0.5)  
34 R: 成人し  
35 ((R: nod))  
36 ((H: gaze in))  
→ 37 R: えーとね  
38 ((R: nod))  
39 ((H: gaze out))  
40 (0.5)  
41 ((R: forward))  
42 (1.5)  
⇒ 43 R: こっち見てよね  
44 ((H: gaze in))  
45 R: 成人式出席後にね  
46 ((R: nod))

### 6.4.3 インタラクション例3

インタラクション例3では Talking-Ally が言い直すと、ユーザの視線を配慮した身体動作とを同時に行った場面を示している。

はじめに Talking-Ally は姿勢を少し後ろに倒すことで、ユーザからの視線を得ている。しかし、しばらくしてユーザの視線が別の方向に向いてしまったため、「もう一回言うよ」とユーザに言い直すことを伝えた。このとき、Talking-Ally は聞き手が向いている方向に視線を向けている。この「もう一回言うよ」という発言と、聞き手の視線を配慮した Talking-Ally の身体動作によりユーザからの視線を得た。さ

らに、それまでよそ見をしていたユーザに対して「こっち見てね」と Talking-Ally を見ることをアピールし、また Talking-Ally の視線を聞き手に向ける身体動作も行うことで、注目をさらに得ようとしている。そうして、聞き手の視線を配慮した身体動作と言い直しを同時に行っている。

以下に着目している場面のトランスクリプトを示す。

- 61 ((H: gaze out))  
62 ((R: back))  
63 ((H: gaze in))  
64 R: 道交法違反の疑いでね  
65 ((R: back))  
66 ((H: gaze out))  
67 ((R: nod))  
68 (1.3)  
69 R: もう一回言うよ  
→70 ((R: Human's gaze))  
71 (0.2)  
72 ((H: gaze in))  
73 (1.5)  
74 R: こっち見てよね  
75 ((R: nod))  
76 R: 道交法違反の疑いでね  
77 ((R: gaze front))  
78 ((R: forward))  
79 ((R: nod))

## 6.5 印象評定

### 6.5.1 実験内容

テキストをそのまま読み上げる Talking-Ally の発話と 6.3 節の方法による Talking-Ally の発話では、Talking-Ally に対する聞き手の構えはどのように異なるのかを調べる実験を行った。

#### 目的

どのような方法で生成された Talking-Ally の発話が聞き手の志向姿勢を引き出すのか調査することを目的とする。

#### 方法

ここでは、発話の生成様式の変化という 1 要因のみの 4 水準で実験を計画した。表 6.3 にその条件を示す。modality や turn-initials の付与の有無は、6.3.2 や表 6.1 で述べたターン開始要素 (表 6.1 中の TI) やモダリティ (表 6.1 中の MD) をニュース文章に付与するかどうかを指している。また、視線への配慮の有無は 6.3.3 や表 6.1 で述べた聞き手の視線に配慮した発話 (表 6.1 中の EB) を Talking-Ally の発話に付与するかどうかを指している。

これらの実験条件に従って、4 つの異なる発話生成システムを作成した。この 4 種類の発話生成システムによる Talking-Ally からの発話を被験者に聞いてもらい、質問紙を用いて印象評価を行なった。なお、被験者は Talking-Ally の発話を録画したビデオを第三者として観察するのではなく、実験室内で実際に Talking-Ally の前に座り、被験者の視線方向の変化に随伴してリアルタイムに産出される Talking-Ally の発話を聞いている。また、被験者の興味を散漫にさせるために、Talking-Ally の後方でビデオ<sup>1</sup>を上映した。

<sup>1</sup>Hanna, W., Barbera, J., Jones, C.: *TOM & JERRY: Sci-Fi & Fantasy type episodes*[DVD], Warner Home Video (2003).

表 6.3: 実験条件

条件	発話の生成様式
A	modality/turn-initials を付与しない 聞き手の視線に配慮しない
B	modality/turn-initials を付与する 聞き手の視線に配慮しない
C	modality/turn-initials を付与しない 聞き手の視線に配慮する
D	modality/turn-initials を付与する 聞き手の視線に配慮する

質問項目を表 6.4 に示す。質問項目の Q1, Q2 は Talking-Ally の発話が理解しやすいものかどうか（了解性），Q3 は Talking-Ally の発話に引き込まれたかどうか（参加態度），Q4–Q10 は聞き手が Talking-Ally をどのような存在として捉えていたか（志向的な構え）をそれぞれ確認するためのものである。

ここで「了解性」と「参加態度」を測定項目として設定した理由は、被験者の聞く姿勢を反映しながら発話を調整した結果、Talking-Ally の発話の内容を被験者はよく理解することができるのか、また、発話者である Talking-Ally に対して、いま Talking-Ally の発話を聞いていますよという仕草を被験者は表出するのかを調べるためである。Talking-Ally の動作に被験者が志向的な構えを採用した場合、被験者から上記のような行動が引き出されると考えた。

それぞれの質問項目は 5 件法を採用し、「よく当てはまる」、「当てはまる」、「どちらともいえない」、「当てはまらない」、「まったく当てはまらない」の 5 段階評価を行なう。

表 6.4: 質問紙

質問項目	質問内容
Q1	ロボットの発話は聞き取りやすかった
Q2	ロボットの発話は理解しやすかった
Q3	ロボットの発話を聞く気になった
Q4	ロボットの発話は独り言のようだった
Q5	ロボットはあなたを聞き手として意識していた
Q6	ロボットに生き物らしさを感じた
Q7	ロボットはあなたに話したがっていると感じた
Q8	ロボットの発話は予め用意されたものだと感じた
Q9	ロボットの発話に伝えたいという意味を感じた
Q10	ロボットは自分の意思で発話・行動していると感じた

## 被験者

被験者は 19 歳から 21 歳の男性 13 名, 女性 1 名の計 14 名である。被験者には「ロボットの会話を 4 パターン聞き, それぞれの会話が終了した後アンケートに答える」, 「後方で流れているビデオの内容も記入する」という 2 つの教示を行った。また, 実験の説明・教示を Talking-Ally に発話させることで, 第 1 施行と第 4 施行での Talking-Ally の音声への違和感や慣れを緩和させた。4 種類の発話方法と発話内容は, カウンターバランスに配慮して被験者に提示した。

表 6.5: 各実験条件における評価値の平均

	条件 A	条件 B	条件 C	条件 D
Q1	<u>2.57</u>	<u>2.57</u>	2.86	<b>3.07</b>
Q2	<u>2.71</u>	2.86	2.93	<b>3.00</b>
Q3	<u>3.57</u>	4.14	3.86	<b>4.15</b>
Q4	<u>2.71</u>	3.14	3.43	<b>3.71</b>
Q5	<u>2.86</u>	3.57	3.43	<b>4.00</b>
Q6	<u>2.86</u>	3.57	3.43	<b>4.07</b>
Q7	<u>3.00</u>	4.21	3.64	<b>4.36</b>
Q8	<u>2.64</u>	3.00	<b>3.21</b>	2.93
Q9	<u>3.07</u>	4.14	3.71	<b>4.21</b>
Q10	<u>2.71</u>	3.36	3.07	<b>3.79</b>

### 6.5.2 結果

各実験条件における評価の平均値を表 6.5 に、標準偏差を表 6.6 に示す。また、これらの結果についてグラフで表したものを、図 6.5、図 6.6 に示す。なお、表 6.5 については各質問項目で最も高い値を太字、低い値を下線で示している。また、評価を分かりやすくするために、ロボットの発話は独り言のようだったかを問う設問 (Q4) とロボットの発話は予め用意されたものと感じたかを問う設問 (Q8) の評価値を反転させている。表 6.5 に示す通り、ロボットの発話は予め用意されたものと感じたかを問う設問 (Q8) を除いた全ての質問項目において条件 D の評価が高いことが分かる。

ここで表 6.5 の結果に対して分散分析 (ANOVA) を用いて実験操作の効果を検定したところ、質問項目 Q3 に対しては有意傾向 ( $p < 0.10$ ) が、Q4, Q5, Q6, Q7, Q9, Q10 に対しては有意差 ( $p < 0.05$ ) が確認できた。この有意傾向と有意差を確認した各質問項目において、2条件ごと (条件 A-B, A-C, A-D, B-C, B-D, C-D という 6つの組合せ) に t 検定を行い、有意差を調べた。t 検定の結果から有意差の



表 6.6: 各実験条件における評価値の標準偏差

	条件 A	条件 B	条件 C	条件 D
Q1	1.089	0.852	1.167	0.829
Q2	0.914	0.663	0.829	0.679
Q3	1.089	0.864	0.770	0.555
Q4	1.139	1.231	0.646	0.994
Q5	1.167	0.938	1.089	0.679
Q6	0.864	1.089	1.016	0.475
Q7	1.240	0.893	1.151	0.633
Q8	1.008	0.961	1.051	1.141
Q9	0.997	0.535	1.069	0.579
Q10	1.139	0.745	1.072	0.699

確認された各組み合わせの検定値について、表 6.7 にまとめる。

次節以降では、これらの結果について「了解性」「参加態度」「志向的な構え」の項目に分けて詳しく述べる。

表 6.7: t 検定の結果 (有意差が現れた項目について)

質問	組合せ	t 値	結果
Q3	(A,B)	$t(13) = 2.2804$	$*p = 0.04 < 0.05$
Q4	(A,D)	$t(13) = 3.0166$	$**p = 0.009 < 0.01$
	(B,D)	$t(13) = 2.5106$	$*p = 0.026 < 0.05$
Q5	(A,B)	$t(13) = 2.3470$	$*p = 0.035 < 0.05$
	(A,D)	$t(13) = 2.7378$	$*p = 0.017 < 0.05$
Q6	(A,B)	$t(13) = 2.9245$	$*p = 0.012 < 0.05$
	(A,D)	$t(13) = 5.0902$	$**p = 0.0002 < 0.01$
	(C,D)	$t(13) = 2.5898$	$*p = 0.022 < 0.05$
Q7	(A,B)	$t(13) = 4.0505$	$**p = 0.0014 < 0.01$
	(A,D)	$t(13) = 3.6462$	$**p = 0.0030 < 0.01$
Q9	(A,B)	$t(13) = 4.8374$	$**p = 0.0003 < 0.01$
	(A,C)	$t(13) = 2.2234$	$*p = 0.045 < 0.05$
	(A,D)	$t(13) = 3.8894$	$**p = 0.002 < 0.01$
	(C,D)	$t(13) = 2.1885$	$*p = 0.047 < 0.05$
Q10	(A,B)	$t(13) = 2.2234$	$*p = 0.045 < 0.05$
	(A,D)	$t(13) = 2.6874$	$*p = 0.019 < 0.05$
	(C,D)	$t(13) = 2.9245$	$*p = 0.012 < 0.05$

表 6.8: Q1 に対する各被験者の評価値

	条件 A	条件 B	条件 C	条件 D
被験者 1	2	2	2	2
被験者 2	2	3	3	3
被験者 3	3	2	4	3
被験者 4	2	4	2	3
被験者 5	2	3	2	4
被験者 6	1	4	2	3
被験者 7	4	3	4	4
被験者 8	2	2	2	2
被験者 9	5	2	4	4
被験者 10	2	2	5	4
被験者 11	4	2	4	4
被験者 12	2	1	1	2
被験者 13	3	3	3	2
被験者 14	2	3	2	3

### 発話の了解性

Talking-Ally の発話を聞き取れ、理解できたかを調査する質問項目 (Q1, Q2) では有意差が現れなかった。音声合成ソフトの Wizard Voice の合成音が明瞭なものでないことがこの要因の一つとして考えられる。

一方で、聞き取りやすさを問う設問 (Q1) に着目し、被験者一人ひとりの評価値を眺めてみると、表 6.8 のように条件に関わりなく全体的に評価が低い被験者と、各条件間で差がある被験者の 2 組が確認できる。後者のような差が確認できる 8 名の被験者 (被験者 3-被験者 7, 被験者 9-被験者 11) の評価値に対して、t 検定を行った。その結果、条件 (B,D) の間に有意差 ( $t(7)=5.612$ ,  $**p = 0.0008 < 0.01$ ) が確認された。このことから、Talking-Ally の発話を聞き取りにくいと感じる被験者がいる一方で、modality/turn-initials を付与し、視線に配慮した発話を行なうことで聞き取りやすいと感じる被験者もいることが分かる。

また、表 6.5, 図 6.5 においても、modality/turn-initials, 視線に配慮した発話の 2 つを取り入れた発話の方が評価値が向上していることが分かる。以上を考慮すると、modality/turn-initials, 視線に配慮した発話の 2 つを考慮した発話は発話の了解性を高めているという捉え方ができる。

### 発話への参加態度

聞き手が Talking-Ally の発話を聞こうとしたかどうかの参加態度を問う設問 (Q3) では条件 (A,B) において有意差がみられた。「聞いてね」や「こっち見てよ」という要求的な発話方略を用いる条件 D ではなく、要求的発話を用いない条件 B に対して有意な差があることから、要求的な発話ではなく、むしろ modality/turn-initials を付与した発話が聞き手の参加態度に影響を与える直接の要因になっているという捉え方ができる。

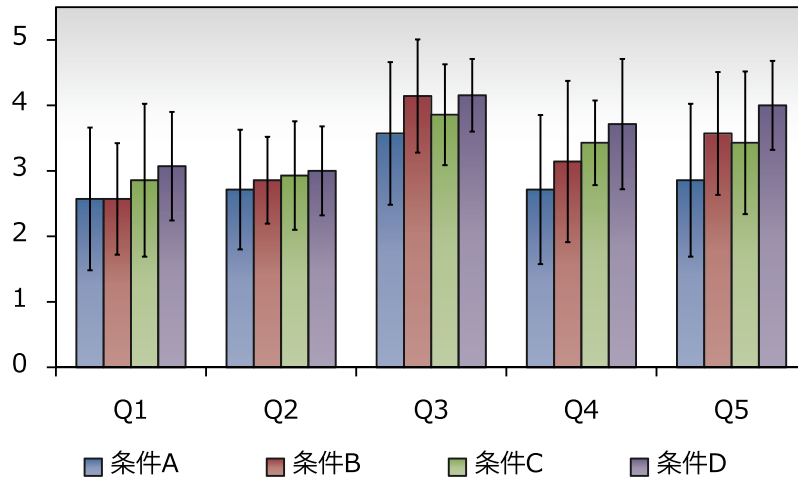


図 6.5: 評価値の平均と標準偏差 (Q1-Q5)

### 聞き手の Talking-Ally に対する構え

聞き手が Talking-Ally をどのような存在として捉えたかを調査する設問 (Q4-Q10) では、ロボットの発話は予め用意されたものと感じたかを問う設問 (Q8) を除いたすべての質問項目において有意差が確認できた。ここから、modality/turn-initials を付与したり、聞き手の視線を配慮した発話を行う方が、聞き手の志向的な構えを引き出しているといえる。

ロボットの発話は予め用意されたものと感じたかを問う設問 (Q8) だけに有意差が現れなかった要因として、次のことが考えられる。被験者が工学部の学生であったために、「ロボットの動作はあらかじめプログラミングされているもの」という先入観が働いたのではないか。また、Web 上のニュース文章を読み上げているという点で、その素材 (記事) があらかじめ用意されていると感じた被験者が多かったと思われる。

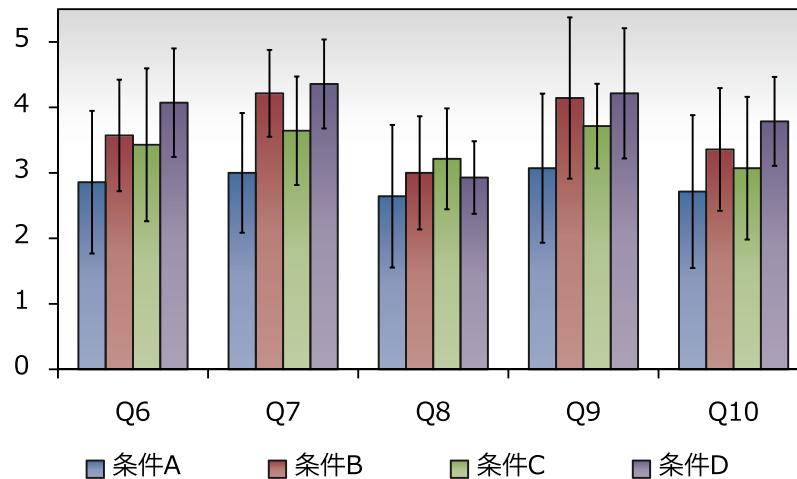


図 6.6: 評価値の平均と標準偏差 (Q6–Q10)

## 6.6 インタラクションログによる被験者の行動分析

前節では、被験者の認知的な側面（ロボットやその会話に対する印象）に与える効果について論じた。つぎに、被験者の身体的な側面に与える効果を論じる。ここで身体的な効果とは、被験者の視線がどのくらい揺り動かされたのかという Talking-Ally の発話の説得性による効果である。実験中に記録した被験者の視線方向の計測データから、Talking-Ally が生成する発話に随伴して、被験者はどのくらい視線方向を調整したのかを探った。

### 6.6.1 被験者の注視領域の推移

分析の第一ステップとして、条件 A と条件 D の 2 条件間での被験者の振る舞いの比較を行った。条件 A の中では turn-initials（以降、TI と略記）や要求的な発

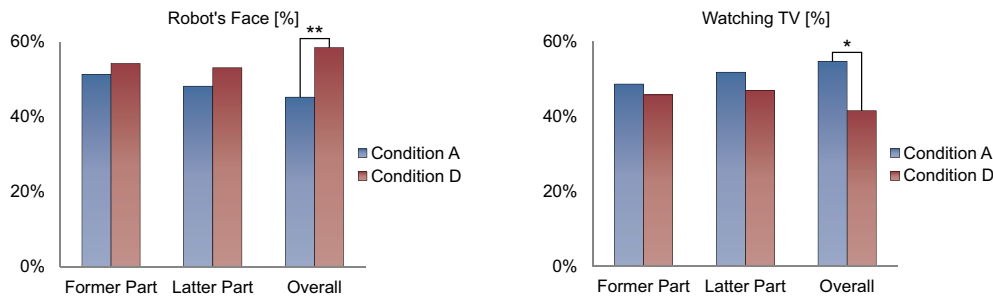


図 6.7: 会話の進行に伴う被験者の視線の変化をグラフで図示したもの

話（以降，EB と略記）は生成されない。一方，条件 *D* では TI や EB は生成される。条件 *D* では，視線の変化を計測し，被験者の関心のある領域（仮想平面）をロボットは推定し，その仮想平面の変化に応じて TI や EB などの発話を適宜行う。TI や EB が生成されることで，被験者の振る舞いにどのような影響が現れるのかを，両条件間の被験者の振る舞いの差異を比較することで確認した。

観測では，60Hz のサンプリングレートで被験者の視線方向の変化を計測した。分析では，被験者の視線が比較的多く得られる次の 2 つの仮想平面を用いた。被験者が TV を見た際に反応する仮想平面（AG5，AG6）とロボットの顔を見た際に反応する仮想平面（AG1，AG2）の 2 種類である。

表 6.9: 会話の進行に伴う被験者の視線の変化（条件 A ⇄ 条件 D）

	ロボットへの視線 (%)		TV への視線 (%)	
	条件 A	条件 D	条件 A	条件 D
会話の前半	51.31%	54.21%	48.69%	45.79%
会話の後半	48.14%	53.03%	51.86%	46.97%
会話全体	45.22%	58.43%	54.78%	41.57%

条件 *A* と条件 *D* におけるロボットへの注視回数と TV への注視回数を算出した（表 6.9）。ここでは，被験者と Talking-Ally とのインタラクションを前半と後半の 2 つに大別している。インタラクションを前半と後半に分ける理由は，インタラク

ションの経過にもなって被験者の関心にどのような変化が現れるのかを探るためである。また、表 6.9 の結果をグラフに可視化した (図 6.7)。

図 6.7 の左に示す通り、条件 *D* では条件 *A* と比較して被験者はより多くの視線をロボットに向けている。これは会話の前半や後半に対して、また、会話全体においても同様である。図 6.7 の左で示される視線データ (全員の被験者) に対して *t* 検定を行ったところ、会話全体のデータにおいて有意傾向を確認することができた ( $t = 4.2015, df = 24, p = 0.0003 < 0.01$ )。ここから、ロボットが TI や EB などの発話を行う条件 *D* の方が被験者の視線をより多く得ていることが示された。もう一つ重要なことは、TI や EB などの発話を伴わない条件 *A* では、時間の経過にもなって、ユーザの視線がロボットから離れてしまう点である。

図 6.7 の右に示す通り、条件 *D* に比べて条件 *A* の方が被験者は TV の方を注視する回数が多い。この図 6.7 の右で示される視線データ (全員の被験者) に対して *t* 検定を実施したところ、会話全体のデータにおいて有意傾向を確認することができた ( $t = 2.1708, df = 24, p = 0.0401 < 0.05$ )。興味深い点は、会話の後半にかけて被験者の TV への注視回数が増加する点である (表 6.9)。ここから、被験者の視線に配慮した TI や EB などの発話をロボットが行わない場合では、時間の経過にもなって被験者の視線が TV の方に移ることが示された。一方、被験者の視線に追従して発話の繰り出される条件 *D* では、多くの被験者が視線をロボットに向けている。以上を考慮すると、被験者のロボットへの注視行動に条件 *D* の生成様式が大きな役割を果たしていることが分かる。

### 6.6.2 ロボットの発話に伴う視線方向の変化

続いて、条件 *D* において、被験者の視線方向はロボットの発話の進行に伴って、どのように時間的に変化していくのかについて調べた。そのためには被験者がいまどの物体に視線を向けているのかを把握する必要がある。ここでは、被験者の視線データ (X, Y 座標の二次元データ) を混合正規分布によりクラスタリングした。被験者の視線の方向は大別して、ロボット、あるいは、TV の方を向いている



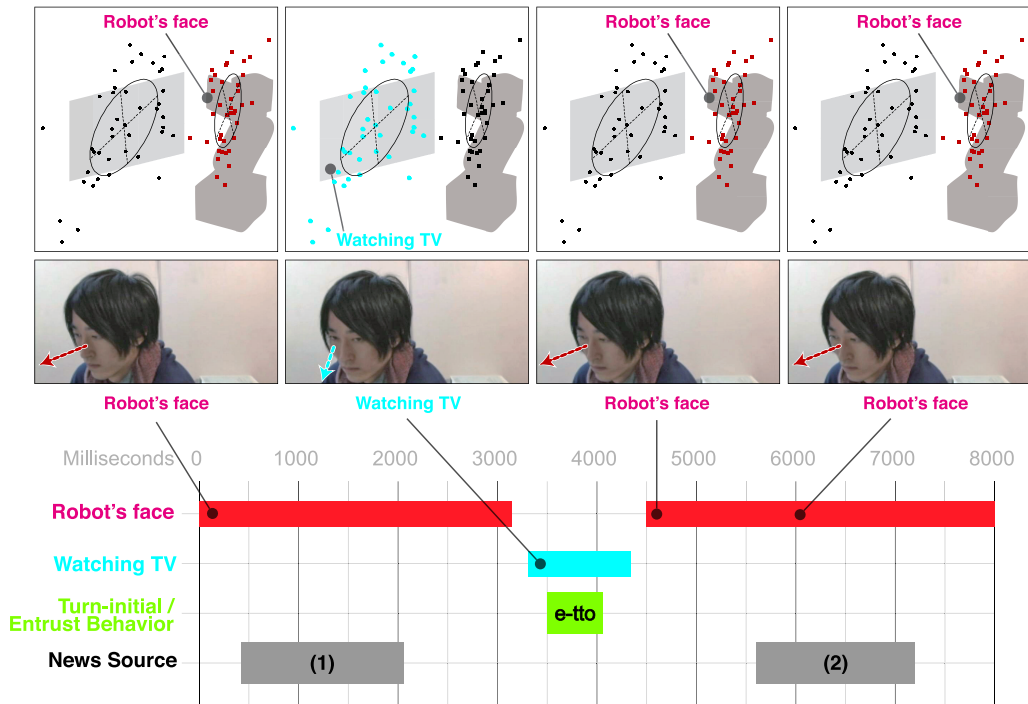


図 6.8: ロボットの会話の進行に伴う被験者の視線方向のバラつき (被験者 A)

ことが予想される。そのため、クラスタリングの要素数をあらかじめ 2 に設定することで 2 つのクラスタを取得した。被験者全員の視線データについてクラスタリングを行った。得られた結果を代表するデータを、被験者の視点で描いたものを図 6.8、図 6.9 の上部に示す。

この図はロボットの発話に対して、被験者の視線がどのように変化するかを図示したものである。横軸は時間軸を示している。ユーザの視線データを 1000 フレームごとに区切り、クラスタリングの結果から、被験者がいまロボットと TV のどちらに視線を向けているのかを算出した。また、それぞれ 2 つのクラスタを円で囲った。

まず、図 6.8 について説明する。ロボットの発話が開始した 0~3000 ミリ秒の間では、被験者の視線はロボットの顔面を向いている。3000 ミリ秒~4000 ミリ秒の間では、被験者は TV の方に視線を向けている。そして 4000 ミリ秒~5000 ミリ秒の間では、ロボットの TI の発話とともに、被験者の視線がロボットの方に

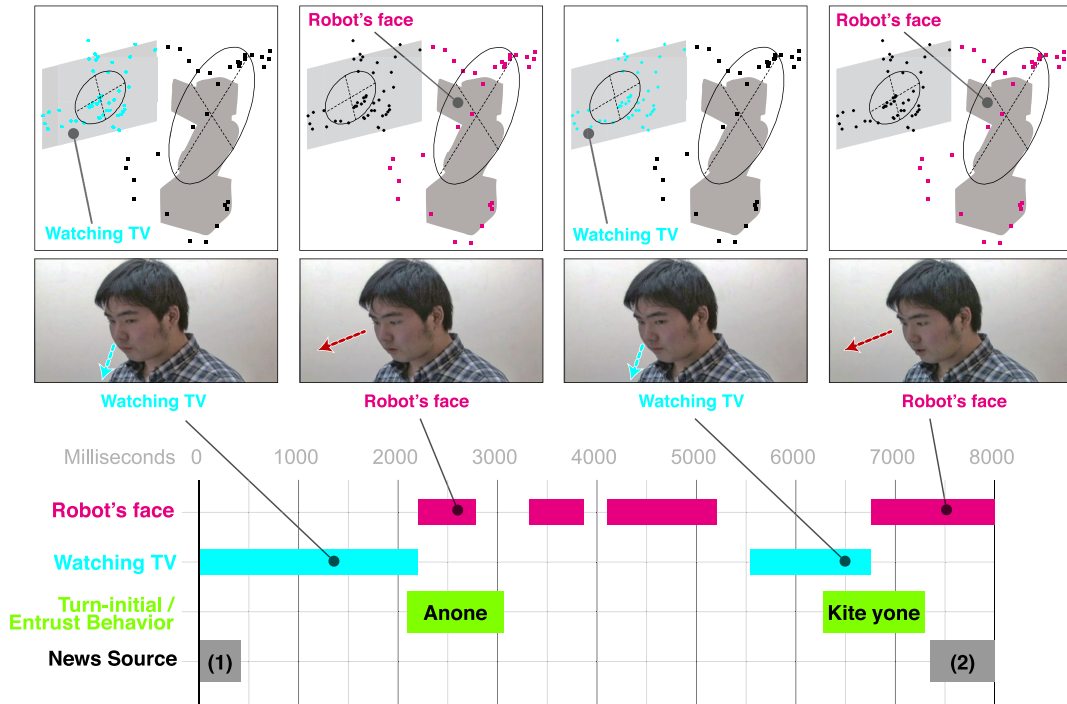


図 6.9: ロボットの会話の進行に伴う被験者の視線方向のバラつき（被験者 B）。

戻っている。

同様の流れを図 6.9 においても確認できる。2000 ミリ秒～3000 ミリ秒の区間において、ロボットは TI の発話を行った。その区間では、被験者の視線方向が変化することはなかった。次に、6000 ミリ秒～8000 ミリ秒の区間で EB の発話を行った。そして、被験者はロボットに視線を向けた。このように、ロボットの TI、あるいは、EB の発話に随伴して、被験者は視線方向を調整していることがわかる。

### 6.6.3 説得性の効果

次に、同じ TI や EB の発話であっても、ロボットがどのような時間的なタイミングで発話するのかの違いによって、被験者の視線方向の調整がどのように変化するのかを調べた。すなわち、条件 B と条件 D の効果の違いを比較した。条件 B では、ロボットは被験者の視線と無関係に、無作為的に TI や EB などの発話を行う。

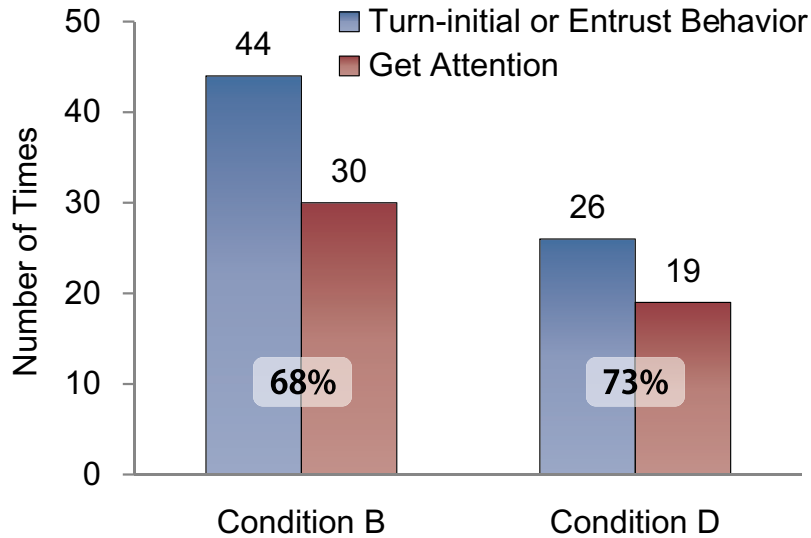


図 6.10: 視線回復の成功率の推移 (条件 B ⇄ 条件 D 間での比較)

被験者の聞き手性の表示を考慮しない条件である。一方、条件 *D* では、ロボットは被験者の視線（聞き手性の表示）を考慮して、TI や EB の発話を行う。この条件 *B* と条件 *D* の効果を比較することで、聞き手性の表示を考慮した場合のロボットの発話（説得性）を見積もる。

両条件間の TI, EB の発話に伴う被験者の視線変化の回数を比べた (図 6.10)。ロボットの TI や EB の発話回数を比較すると、条件 *B* の方が多い。しかし、ロボットの TI, EB の発話に対して、被験者がどのくらい視線を回復したのか（確率）は条件 *D* の方が高い。条件 *B* は 68% であり、条件 *D* は 73% である。

#### 6.6.4 反応時間の評価

説得性の効果を確かめるもう一つの指標として、被験者の視線回復に要する反応時間を調べた。ロボットの説得性の効果が弱ければ、被験者の行動を変化させ

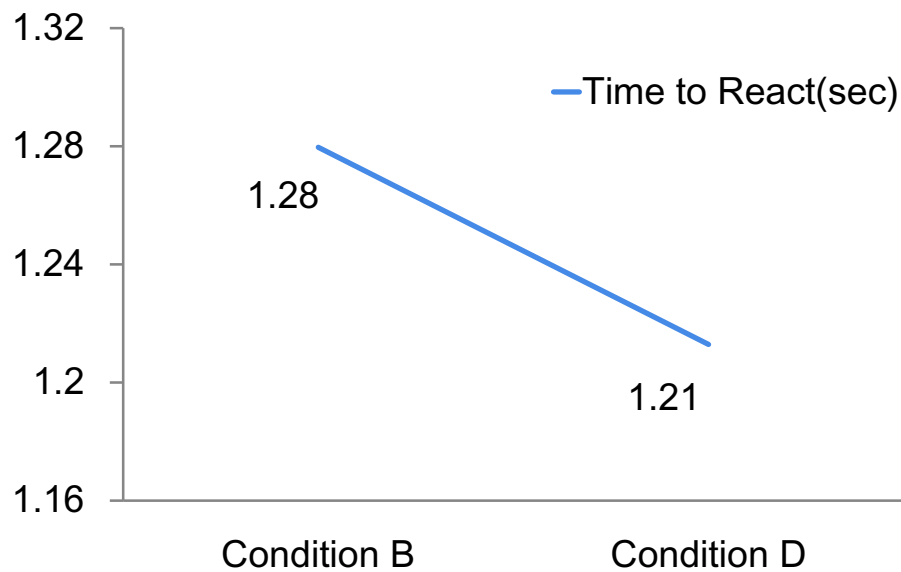


図 6.11: 視線回復に要する時間 [秒] の推移 (条件 B ⇄ 条件 D 間での比較)

るためにはより多くの時間を費やす。それはロボットのコミュニケーションに被験者はどのくらい理解できたのかという理解度の観点と、ロボットのコミュニケーションが被験者に与えた影響力という二つの観点から整理することができる [51]。

会話の半ばにおけるロボットの TI や EB の発話に対して、代表的な被験者が視線回復に要した反応時間を図 6.11 と図 6.12 に図示した。条件 B と条件 D では傾向が異なる。図 6.11 は条件 B よりも条件 D の方が反応時間が短いことを示している。従って、聞き手性の表示と宛名性の二つを考慮した場合のほうが、より素早く被験者は視線方向を変化させていることが分かる。

図 6.12 では次のことが分かる。条件 B では、時間の経過とともに、被験者の反応時間は急激に上昇している。一方、条件 D では、下降し始めている。同様の傾向が別の被験者においても確認された。

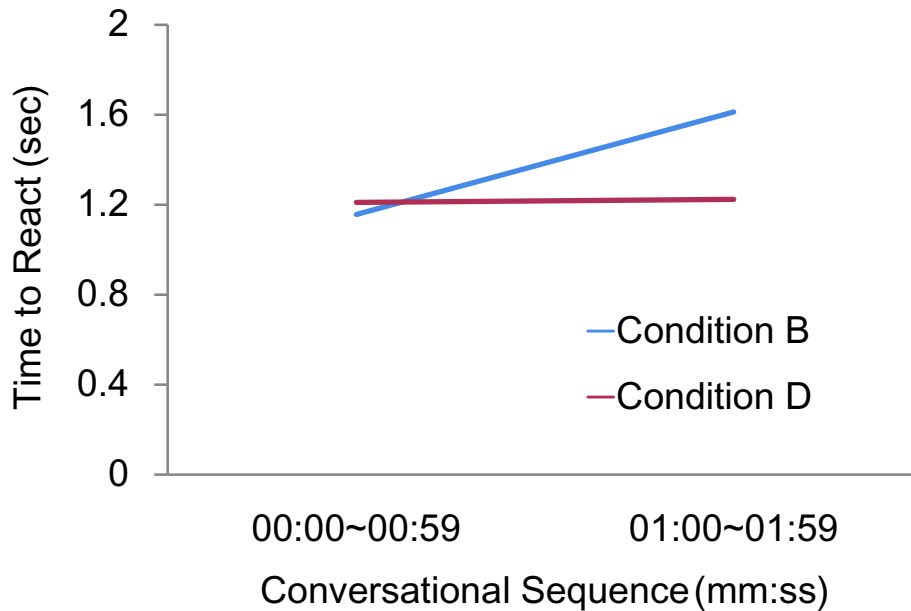


図 6.12: 視線回復に要する時間 [秒] の推移 (会話の前半⇄後半での比較)

### 6.6.5 被験者—ロボットの相互に与えた影響

ロボットの動作と被験者の動作がインタラクションの流れの中でどのように相互に影響していたのかを探ることは、説得性の効果を探るうえで非常に有用である。ここでは、ロボットの反応時間と被験者の反応時間の二つを比較した。上記の効果を探る目的で条件 D の視線データとロボットの動作ログを評価した。条件 D は、ロボットは被験者の聞き手性の表示を考慮し、発話の宛名性を向上させる条件である。前項までの分析で示したように、同条件では顕著な効果が現れている。

ここでは、ロボットが TI や EB の発話を行った際の被験者の反応時間 (図 6.13, 図 6.14, 図 6.15 の青線) と被験者の視線がロボットから外れた際にロボットが TI や EB の発話を行うまでの時間 (図 6.13, 図 6.14, 図 6.15 の赤線) の二つを図示している。これら 2 つの反応時間の相互の関わりが図 6.13, 図 6.14, 図 6.15 に示されている。

ここではインタラクションの経過を 3 つに大別し、それぞれの時間的な区切り

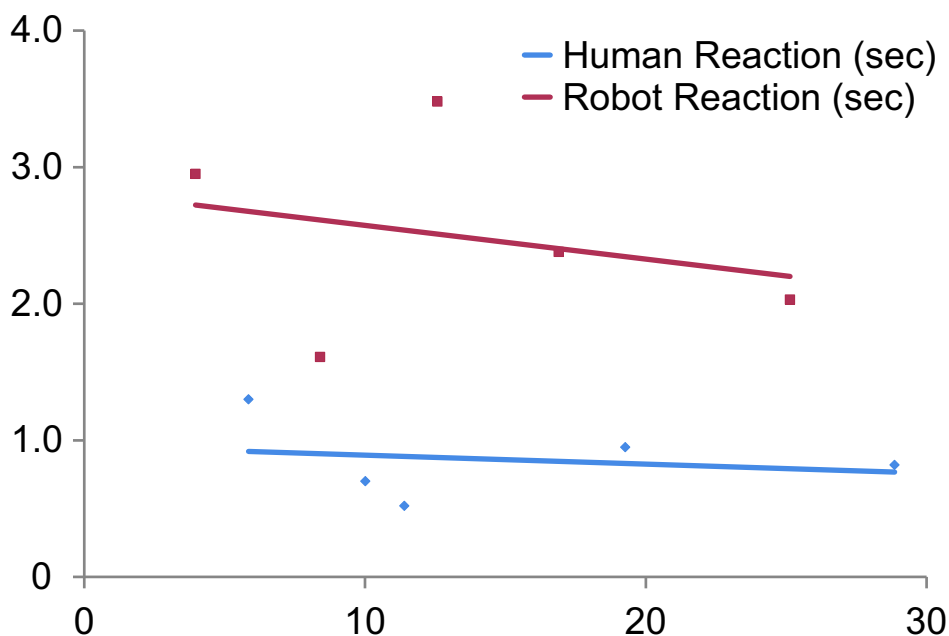


図 6.13: 被験者の反応時間とロボットの反応時間 (セグメンテーション 1)

(セグメンテーションと呼ぶ) の中で、どのような傾向があるのかを確かめた。そのために、各セグメンテーションにおける2種類の反応時間に対して、一次の線形近似を行った。一次の線形近似を用いることで、データの傾向をおおまかに確認できることが一般に知られている [44]。

1つ目のセグメンテーション (図 6.13) では、被験者の反応時間とロボットの反応時間の間には大きなギャップが存在する。しかし、2つ目のセグメンテーション (図 6.14) では、被験者はロボットと同程度の反応時間を保つようになった。そして、3つ目のセグメンテーション (図 6.15) では、ロボットの反応時間と被験者の反応時間の2つが同程度のまま、より高速になっていることが分かる。

ここではロボットに学習のメカニズムを組み入れていない。しかし、聞き手の振る舞いとロボットの発話が相互に影響しあうようにインタラクションをデザインした。従って、聞き手性の表示 (被験者の振る舞い) を反映したロボットの発話に被験者が適応的に振る舞うことで、あたかもそのロボットが適応的に振舞っているように見える。このような理由から、被験者の振る舞いが素早くなること

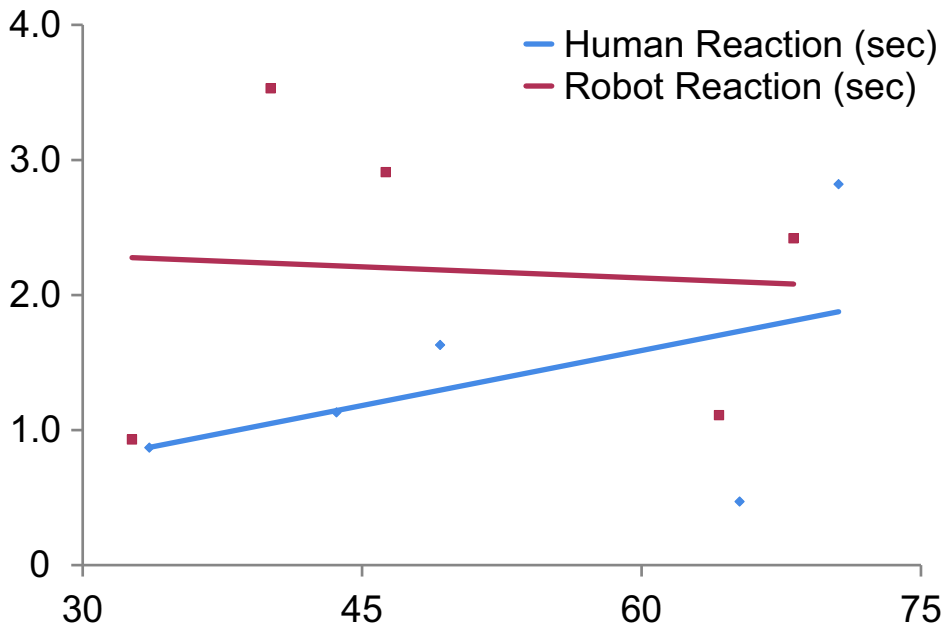


図 6.14: 被験者の反応時間とロボットの反応時間 (セグメンテーション 2)

で、ロボットの動作も素早くなるという結果 (図 6.15) が得られた。

## 6.7 総括と今後の展望

本研究では、聞き手に対する宛名性や聞き手との相互行為的な調整を含む発話生成システムの提案を行い (6.2 節)、ロボットへの実装を行った (6.3 節, 6.4 節)。テキストを淡々と読み上げる従来の音声合成システムと比較して、提案システムは人の志向的な構えを引き出す傾向があることが分かった (6.5 節)。また、被験者の視線方向の変化を手がかりに、ロボットの発話における説得性の効果 (=他者を揺り動かすような力) を探った (6.6 節)。次の 2 点が 6.5 節や 6.6 節の実験結果から明らかになった。

一つは、6.5.2 の「聞き手の Talking-Ally に対する構え」について、modality/turn-initials を踏まえた発話とユーザの視線を配慮した発話の両方を用いることで、聞き手は Talking-Ally から志向性を感じる傾向がある。そして、このような傾向は

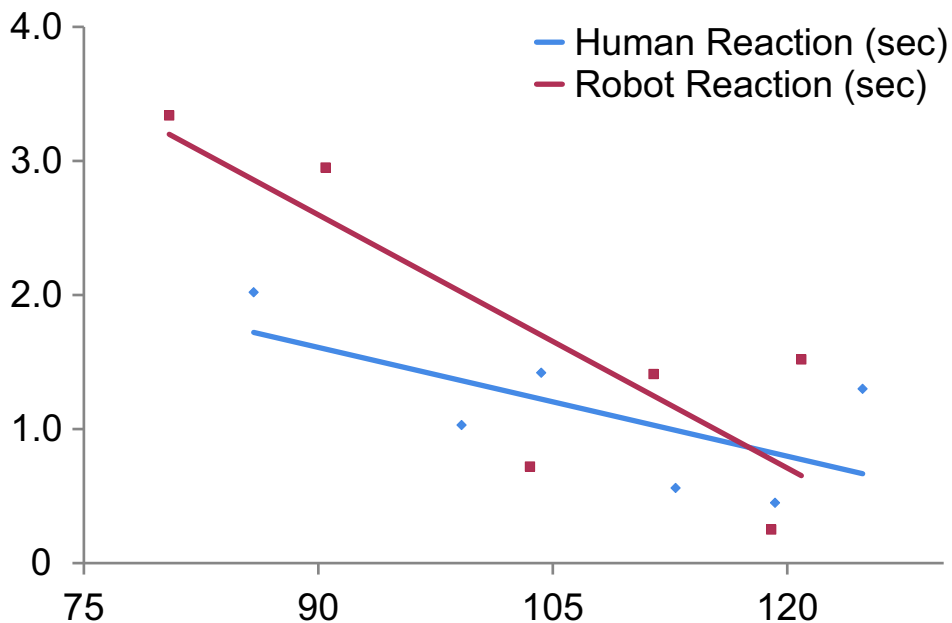


図 6.15: 被験者の反応時間とロボットの反応時間 (セグメンテーション 3)

印象評定の分析のみによるものではなく、実測データを用いた分析においても示されている。まず A 条件と D 条件を比較すると、D 条件の方が高い。これは印象評定の結果が示すとおり理由である。次に、B 条件と D 条件を比較すると、D 条件の方が高い。こちらについても、印象評定の結果と同様である。このように、単に modality/turn-initials を発話するのではその効力は時間の経過とともに次第に薄れていくが、ユーザの視線を配慮して modality/turn-initials を繰り返すことで Talking-Ally の発話に志向性を感じる、発話の説得性が高められていくことが分かった。

もう一つは、実験を行ったことで、新たな課題が見つかった。合成音声に対する個人の好みや聞き手の予備知識など、個人差によって結果が大きく異なる。まず、6.5.2 の発話を理解できたかどうかの「発話の了解性」については、ソフトウェア構成という側面から再考の余地が残る。これは、システムで使用した音声合成ソフトが子どものように可愛らしい声を発話するという特徴を備えていると同時に、その合成音声を初めて聞いた人に対しては聞き取りにくいといった欠点を



---

もつことがその要因の一つになっていると考えられる。今後は、声質などの特徴を変化させて同様の実験を行いたい。また、6.5.2の「発話への参加態度」については、被験者人数の増加や被験者集団の統制を行い、より精確な結果を得る。



## 第7章 結論

本論文では、ユーザーの行動を思わず揺り動かすことの可能な説得性を伴う、新たなメディア技術、コミュニケーション技術について論じた。これまでパースエイシブ・ロボティクスの多くは、ユーザーからの志向的な構えを引き出すエージェントやロボットのデザイン手法、多様なモダリティを付与することでコミュニケーションにおけるリアリティを高める手法などに関心が向けられてきた。そこで、本研究では、会話参与者間の共同行為を引き出す場を媒介として、他者の行動を揺り動かす説得性が社会的に構成されることを指摘した。さらに、本提案にある新たな観点に基づき、多様なインタラクティブ・メディアや発話生成を行うロボットを設計・構築し、それらがユーザーの行動を思わず揺り動かす説得性を有することを、フィールド実験結果に対するエスノメソドロジ的な分析を通して詳細に検証した。またこれらの知見を体系付けることで、インタラクティブ・メディアやソーシャルなロボットの設計指針を整理した。

具体的には、第1章において、本研究と従来研究との差異を明らかにし、本研究の立ち位置を明確にした。また、本研究の中心となるアイデアを整理し、基本的なコンセプトについて述べた。参与者同士のやりとりの中で構成される説得性（＝説得性の社会的構成）について着目した（1.3節）。第2章では、本研究を推進するための初期的なステップとして、第1章のOptic Flow Fieldに関する議論を、人と人の社会的なインタラクションに展開した（＝Table Talk Plus）。人と人との社会的なインタラクションを媒介するインタラクティブメディア（クリーチャ）の移動速度が第三者（実験者）によって操作されることで、発話調整というレベルにおいて、人の身体が揺り動かされることを明らかにした。第3章では、Table Talk Plusと同様のシステムを、ネットワークを介した遠隔地間において実現する

ことで、Table Talk Plus における人の揺動運動が引き起こされる根本的なメカニズムを浮き彫りにした (= Nexus+)。ここでは、相手との関係を形成・維持しながら相手を“特定”する行為がキーになっていると指摘した。第4章では、Nexus+の議論を受けて、会話の流れに合わせながら発話の少ない参加者をスポットライトで“特定”するインタラクティブメディア (= Sociable Spotlight) を新たに構築し、Nexus+において着目した理論を強化を支えるエビデンスを収集した。評価実験の結果、複数人によるコミュニケーションの「場」を揺り動かす効果が確認できた。以上に挙げた Table Talk Plus, Nexus+, Sociable Spotlight という実システムを用いた実験結果に基づく分析によって、第1章の議論が裏付けられた。これらの実験結果から得られた知見は HRI/HAI への応用・展開に重要である。

本研究では、人とロボット/ソフトウェアエージェントとのインタラクションにおける応用例を示すことも、本研究の目的を実現するうえで重要な役割を果たすものであるという指摘を行った。そこで、第5章では、CGアニメーションによってデザインされたソフトウェアエージェントを新たに構築することで、HAIにおける応用例を示した。また、第6章では、HRIにおける応用例を示すために新たな発話生成システム (= Talking-Ally) を構築した。この Talking-Ally を用いた実験では、同様の Turn-initial (例えば「あのね」という言葉) であっても、ユーザの視線方向と無関係に発話される場合 (=B 条件) とユーザの状態に配慮して発話される場合 (=D 条件) という別々の発話方法では、時間の経過とともにその言葉の説得力が低下してしまう (B 条件) か、一定のまま維持される (D 条件) という違いがあることを明らかにした。これは本研究のキー概念である「社会的に構成される説得力 (D 条件)」の効果を HRI の水準において実現したものであり、本研究の目的が達成されたことを実験的に実証している。一方で、本研究の限界も見つかっている。工学の分野では比較的新しい分野にある HRI/HAI では、その実験手法や評価/分析手法自体がまだ十分に確立されていないという課題がある。本研究で触れた実験手法やデータ分析/評価手法についても、現段階における最適な手法を適用したものであり、一般性や再現性という面において、今後議論する余地が残されている。しかし、人とロボットのリアリティを伴ったコミュニケーション

---

ンを可能にする技術について、説得性の社会的構成に着目することで、より身体レベルでの行動において人の身体を揺り動かすことができるということを確認した。このような着眼点・手法はHRI/HAIの研究分野において、未開拓の領域であり、今後はこの「説得性の社会的構成」によるHRI/HAIを広く社会に波及し、一般のユーザ/消費者に対してどのような貢献ができるのか、その実験/評価手法も併せて探っていく。



## 参考文献

- [1] Henny Admoni, Bradley Hayes, David Feil-Seifer, Daniel Ullman, and Brian Scassellati. Are You Looking at Me?: Perception of Robot Attention is Mediated by Gaze Type and Group Size. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, pp. 389–396. IEEE Press, 2013.
- [2] Sean Andrist, Erin Spannan, and Bilge Mutlu. Rhetorical Robots: Making Robots More Effective Speakers Using Linguistic Cues of Expertise. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, pp. 341–348. IEEE Press, 2013.
- [3] Emilia I Barakova and Tino Lourens. Expressing and Interpreting Emotional Movements in Social Games with Robots. *Personal and Ubiquitous Computing*, Vol. 14, No. 5, pp. 457–467, 2010.
- [4] John N. Bassili. Temporal and Spatial Contingencies in the Perception of Social Events. *Journal of Personality and Social Psychology*, Vol. 33, No. 6, pp. 680–685, 1976.
- [5] Cynthia Breazeal. *Designing Sociable Robots*. MIT Press, 2002.
- [6] Cynthia Breazeal. Toward Sociable Robots. *Robotics and Autonomous Systems*, Vol. 42, No. 3-4, pp. 167–175, 2003.

- 
- [7] Crystal Chao and Andrea L. Thomaz. Turn Taking for Human-Robot Interaction. In *Dialog with Robots: Papers from the AAAI Fall Symposium*, pp. 132–134, 2010.
- [8] Vijay Chidambaram, Yueh-Hsuan Chiang, and Bilge Mutlu. Designing Persuasive Robots: How Robots Might Persuade People Using Vocal and Non-verbal Cues. In *HRI*, pp. 293–300, 2012.
- [9] Herbert H. Clark and Thomas B. Carlson. Hearers and Speech Acts. *Language*, pp. 332–373, 1982.
- [10] James W Davis and Aaron F Bobick. Virtual PAT: A Virtual Personal Aerobics Trainer. In *Workshop on Perceptual User Interfaces*, pp. 13–18, 1998.
- [11] Daniel C. Dennett. *The Intentional Stance*. MIT Press, 1989. (若島正, 河田学 訳 (1996) 『志向姿勢の哲学』 白揚社).
- [12] Daniel C. Dennett. *Kinds of minds: Towards an understanding of consciousness*. Phoenix, 1996.
- [13] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. Legibility and Predictability of Robot Motion. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*, pp. 301–308. IEEE, 2013.
- [14] S. W. Draper. Display Managers as a Basis for Usermachine Interaction. *Lawrence Erlbaum Associates*, 1986.
- [15] Starkey Duncan. Some Signals and Rules for Taking Speaking Turns in Conversations. *Journal of Personality and Social Psychology*, Vol. 23, No. 2, p. 283, 1972.



- 
- [16] Derek Edwards and David Middleton. Joint Remembering: Constructing an Account of Shared Experience through Conversational Discourse. *Discourse processes*, Vol. 9, No. 4, pp. 423–459, 1986.
- [17] Nicolas Fay, Simon Garrod, and Jean Carletta. Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size. *Psychological Science*, Vol. 11, No. 6, pp. 481–486, 2000.
- [18] David Feil-Seifer and Maja J Mataric. Defining socially assistive robotics. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pp. 465–468. IEEE, 2005.
- [19] François Ferland, Arnaud Aumont, Dominic Létourneau, and François Michaud. Taking Your Robot for a Walk: Force-guiding a Mobile Robot Using Compliant Arms. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, pp. 309–316. IEEE Press, 2013.
- [20] B. J. Fogg. Captology: The Study of Computers As Persuasive Technologies. In *CHI '97 Extended Abstracts on Human Factors in Computing Systems, CHI EA '97*, pp. 129–129, New York, NY, USA, 1997. ACM.
- [21] B. J. Fogg. Persuasive Computers: Perspectives and Research Directions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '98*, pp. 225–232, 1998.
- [22] B. J. Fogg. *Persuasive Technology: Using Computers to Change what we Think and Do*. Morgan Kaufmann, Amsterdam, 2003.
- [23] Gainer Book Labo, ぐるぐる研究室. + *GAINER*. O'Reilly, 2008.
- [24] J. J. Gibson. 生態学的視覚論: ヒトの知覚世界を探る (古崎敬・古崎愛子・辻敬一郎・村瀬 旻, 訳), 1985.

- 
- [25] Brian Gleeson, Karon MacLean, Amir Haddadi, Elizabeth Croft, and Javier Alcazar. Gestures for Industry: Intuitive Human-robot Communication from Human Observation. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, pp. 349–356. IEEE Press, 2013.
- [26] Erving Goffman. Footing. *Semiotica*, Vol. 25, pp. 1–29, 1979.
- [27] Erving Goffman. *Forms of Talk*. Univ of Pennsylvania Press, 1981.
- [28] Charles Goodwin. The Interactive Construction of a Sentence in Natural Conversation. In *Everyday Language: Studies in Ethnomethodology*, pp. 97–121. Publishers, Inc, 1979.
- [29] Charles Goodwin. Academic Press, 1981.
- [30] Jorne Grolleman, Betsy van Dijk, Anton Nijholt, and Andrée van Emst. Break the Habit! Designing an E-therapy Intervention Using a Virtual Coach in Aid of Smoking Cessation. In *Persuasive Technology*, pp. 133–141. Springer, 2006.
- [31] J. R. C. Ham and C. J. H. Midden. Persuasive Robots Can Help Save Energy: The Influence of the Agency of a Robot on its Power to Influence Sustainable Behavior. *The 9th International Conference on Environmental Psychology 2011*, 2011.
- [32] Jaap Ham, René Bokhorst, Raymond H. Cuijpers, David van der Pol, and John-John Cabibihan. Making Robots Persuasive: The Influence of Combining Persuasive Strategies (Gazing and Gestures) by a Storytelling Robot on Its Persuasive Power. In *ICSR*, pp. 71–83, 2011.
- [33] Fritz Heider and Marianne Simmel. An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, Vol. 57, No. 2, pp. 243–259, 1944.

- 
- [34] Guy Hoffman and Keinan Vanunu. Effects of robotic companionship on music enjoyment and agent perception. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*, pp. 317–324. IEEE, 2013.
- [35] Wijnand IJsselsteijn, Yvonne de Kort, Cees J. H. Midden, Berry Eggen, and Elise van den Hoven, editors. *Persuasive Technology, First International Conference on Persuasive Technology for Human Well-Being, PERSUASIVE 2006, Eindhoven, The Netherlands, May 18-19, 2006, Proceedings*, Vol. 3962 of *Lecture Notes in Computer Science*. Springer, 2006.
- [36] Wijnand IJsselsteijn, Yvonne de Kort, Joyce Westerink, Marko de Jager, and Ronald Bonants. Fun and Sports: Enhancing the Home Fitness Experience. In *Entertainment Computing-ICEC 2004*, pp. 46–56. Springer, 2004.
- [37] 石黒浩, 宮下敬宏, 神田崇行. 知の科学コミュニケーションロボット. 知の科学. オーム社, 2005.
- [38] Hiroshi Ishii, Craig Wisneski, Julian Orbanes, Ben Chun, and Joe Paradiso. PingPongPlus: Design of an Athletic-tangible Interface for Computer-supported Cooperative Play. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99*, pp. 394–401, New York, NY, USA, 1999. ACM.
- [39] Cory D Kidd and Cynthia Breazeal. Effect of a Robot on User Perceptions. In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, Vol. 4, pp. 3559–3564. IEEE, 2004.
- [40] Hiroyuki Kidokoro, Takayuki Kanda, Drazen Brcic, and Masahiro Shiomi. Will I Bother Here?—A Robot Anticipating its Influence on Pedestrian Walking Comfort. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*, pp. 259–266. IEEE, 2013.

- 
- [41] Kazuki Kobayashi and Seiji Yamada. Making a Mobile Robot to Express its Mind by Motion Overlap. *Advances in Human-Robot Interaction*, pp. 111–124, 2009.
- [42] H. Kose-Bagci, K. Dautenhahn, and C. L. Nehaniv. Emergent Dynamics of Turn-taking Interaction in Drumming Games with a Humanoid Robot. In *17th IEEE Int. Symp. on Robot and Human Interactive Communication*, pp. 346–353. IEEE, 2008.
- [43] Peter Noble Kugler and Michael T Turvey. *Information, Natural Law, and the Self-assembly of Rhythmic Movement*. Lawrence Erlbaum Associates, Inc, 1987.
- [44] Arno B. J. Kuijlaars. Book Review: "A Course in Approximation Theory" by Ward Cheney and Will Light. *Journal of Approximation Theory*, Vol. 112, No. 2, p. 318, 2001.
- [45] Yoshinori Kuno, Kazuhisa Sadazuka, Michie Kawashima, Keiichi Yamazaki, Akiko Yamazaki, and Hideaki Kuzuoka. Museum Guide Robot Based on Sociological Interaction Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1191–1194. ACM, 2007.
- [46] Hideaki Kuzuoka, Yuya Suzuki, Jun Yamashita, and Keiichi Yamazaki. Reconfiguring Spatial Formation Arrangement by Robot Body Orientation. In *Proceedings of the 5th ACM/IEEE International Conference on Human-robot Interaction*, pp. 285–292. IEEE Press, 2010.
- [47] Denis N Lee and H Kalmus. The Optic Flow Field: The Foundation of Vision [and Discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, Vol. 290, No. 1038, pp. 169–179, 1980.

- 
- [48] DN Lee and JR Lishman. Visual Proprioceptive Control of Stance. *Journal of Human Movement Studies*, 1975.
- [49] Willem JM Levelt. *Speaking: From intention to articulation*, Vol. 1. MIT press, 1993.
- [50] Andrés Lucero, Rodrigo Zuloaga, Selene Mota, and Felipe Muñoz. Persuasive Technologies in Education: Improving Motivation to Read and Write for Children. In *Persuasive Technology*, pp. 142–153. Springer, 2006.
- [51] R.J. Lutz and J. L. Swasy. *Integrating Cognitive Structure And Cognitive Response Approaches To Monitoring Communications Effects*. Association for Consumer Research, 1977.
- [52] バフチン M.M. 小説の言葉, 1979.
- [53] Bronislaw Malinowski. The Problem of Meaning in Primitive Languages. *Language and Literacy in Social Practice: A Reader*, pp. 1–10, 1994.
- [54] Teddy McCalley, Florian Kaiser, Cees Midden, Merijn Keser, and Maarten Teunissen. Persuasive Appliances: Goal Priming and Behavioral Response to Product-integrated Energy Feedback. In *Persuasive Technology*, pp. 45–49. Springer, 2006.
- [55] Margaret E Morris. Social Networks as Health Feedback Displays. *Internet Computing, IEEE*, Vol. 9, No. 5, pp. 29–37, 2005.
- [56] Jonathan Mumm and Bilge Mutlu. Human-robot Proxemics: Physical and Psychological Distancing in Human-robot Interaction. In *Proceedings of the 6th International Conference on Human-robot Interaction*, pp. 331–338. ACM, 2011.

- 
- [57] Bilge Mutlu, Toshiyuki Shiwa, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. Footing in Human-robot Conversations: How Robots Might Shape Participant Roles Using Gaze Cues. In *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, HRI2009*, pp. 61–68, 2009.
- [58] Kayako Nakagawa, Masahiro Shiomi, Kazuhiko Shinozawa, Reo Matsumura, Hiroshi Ishiguro, and Norihiro Hagita. Effect of Robot’s Active Touch on People’s Motivation. In *Proceedings of the 6th International Conference on Human-robot Interaction*, pp. 465–472. ACM, 2011.
- [59] Stefanos Nikolaidis and Julie Shah. Human-robot Cross-training: Computational Formulation, Modeling and Evaluation of a Human Team Training Strategy. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, pp. 33–40. IEEE Press, 2013.
- [60] Donald A. Norman. *User Centered System Design: New Perspectives in Humancomputer Interaction*. Lawrence Erlbaum Associates, 1986.
- [61] Donald A. Norman. *Cognitive Artifacts*. Cambridge University Press, 1991.
- [62] Masa Ogata, Yuta Sugiura, Hirotaka Osawa, and Michita Imai. Pygmy: A Ring-shaped Robotic Device that Promotes the Presence of an Agent on Human Hand. In *Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction*, pp. 85–92. ACM, 2012.
- [63] Ryota Oiwa, Mitsuyo Hashida, and Haruhiro Katayose. Crossing Colorful Communication: A System for Enhancing Sound Communication. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology, ACE '07*, pp. 296–297, New York, NY, USA, 2007. ACM.

- 
- [64] Michio Okada, Shoji Sakamoto, and Noriko Suzuki. Muu: Artificial Creatures as an Embodied Interface. In *27th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH 2000), The Emerging Technologies: Point of Departure*, 2000.
- [65] Tyler Pace, Shruti Ramalingam, and David Roedl. Celerometer and Idling Reminder: Persuasive Technology for School Bus Eco-driving. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems, CHI EA '07*, pp. 2085–2090, New York, NY, USA, 2007. ACM.
- [66] Aaron Powers, Sara Kiesler, Susan Fussell, and Cristen Torrey. Comparing a Computer Agent with a Humanoid Robot. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*, pp. 145–152. IEEE, 2007.
- [67] Akanksha Prakash, Jenay M Beer, Travis Deyle, Cory-Ann Smarr, Tiffany L Chen, Tracy L Mitzner, Charles C Kemp, and Wendy A Rogers. Older Adults' Medication Management in the Home: How Can Robots Help? In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*, pp. 283–290. IEEE, 2013.
- [68] E. S. Reed. *アフオーダンスの心理学：生態心理学への道*. 新曜社, 2000.
- [69] Byron Reeves and Clifford Nass. *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press, 1996.
- [70] Byron Reeves and Clifford Nass. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, 1996. (細馬宏通 訳 (2001) 『人はなぜコンピューターを人間として扱うか-「メディアの等式」の心理学』翔泳社).

- 
- [71] C. Reynolds. Flocks, Herds, and Schools: A Distributed Behavioral Model. *Computer Graphics*, Vol. 21, No. 4, pp. 25–34, 1987.
- [72] Maïke A. J. Roubroeks, Jaap Ham, and Cees J. H. Midden. When Artificial Social Agents Try to Persuade People: The Role of Social Agency on the Occurrence of Psychological Reactance. *I. J. Social Robotics*, Vol. 3, No. 2, pp. 155–165, 2011.
- [73] Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn-taking for conversation. *Language*, pp. 696–735, 1974.
- [74] Kazuhiko Shinozawa, Futoshi Naya, Junji Yamato, and Kiyoshi Kogure. Differences in Effect of Robot and Screen Agent Recommendations on Human Decision-making. *International Journal of Human-Computer Studies*, Vol. 62, No. 2, pp. 267–279, 2005.
- [75] Masahiro Shiomi, Takayuki Kanda, Satoshi Koizumi, Hiroshi Ishiguro, and Norihiro Hagita. Group Attention Control for Communication Robots with Wizard of OZ Approach. In *Human-Robot Interaction (HRI), 2007 2nd ACM/IEEE International Conference on*, pp. 121–128. IEEE, 2007.
- [76] Mikey Siegel, Cynthia Breazeal, and Michael I. Norton. Persuasive Robotics: The Influence of Robot Gender on Human Behavior. In *IROS*, pp. 2563–2568, 2009.
- [77] Herbert Alexander Simon. *The sciences of the artificial*. MIT press, 1996.
- [78] Noriko Suzuki, Yugo Takeuchi, Kazuo Ishii, and Michio Okada. Evaluation of Affiliation in Interaction with Autonomous Creatures. In *EuroSpeech*, 1999.



- 
- [79] Yasutaka Takeda, Taisuke Miyake, Hironori Uto, Yuta Yoshiike, R Ravindra De Silva, and Michio Okada. COLUMN: A Novel Architecture for Transformable Artifact. In *Proc. of VRIC*, pp. 271–277, 2010.
- [80] Deborah Tannen. *You Just Don't Understand: Women and Men in Conversation*. William Morrow Paperbacks, 2001.
- [81] Andrea L. Thomaz and Crystal Chao. Turn-taking Based on Information Flow for Fluent Human-robot Interaction. *AI Magazine*, Vol. 32, No. 4, pp. 53–63, 2011.
- [82] Cristen Torrey, Susan R Fussell, and Sara Kiesler. How a Robot Should Give Advice. In *Human-Robot Interaction (HRI), 2013 8th ACM/IEEE International Conference on*, pp. 275–282. IEEE, 2013.
- [83] Yuki Ueno, Hideo Kitagawa, Kiyooki Kakihara, and Kazuhiko Terashima. Design and Control for Collision Avoidance of Power-assisted Omnidirectional Mobile Wheelchair System. In *System Integration (SII), 2011 IEEE/SICE International Symposium on*, pp. 902–907. IEEE, 2011.
- [84] Andrew Wang. Physically Animated Desktop Computer for Ergonomic & Affective Movement. Master's thesis, 2006.
- [85] 山崎敬一, 西阪仰. 語る身体・見る身体. ハーベスト社, 1997.
- [86] Yuta Yoshiike, P. Ravindra De Silva, and Michio Okada. Cues for Social PC: Coordinate and Synchronize its Cues Based on User Attention and Activities on Display. In *HRI*, pp. 135–136, 2010.
- [87] Kuanhao Zheng, Dylan F Glas, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. Supervisory Control of Multiple Social Robots for Navigation. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, pp. 17–24. IEEE Press, 2013.

- [88] ノイマンピアノ. 2061: Max オデッセイ: 音楽と映像をダイナミックに創造する! 最高の開発環境を徹底解説. リットーミュージック, 2006.
- [89] ジェームス V. ワーチ. 心の声: 媒介された行為への社会文化的アプローチ. 福村出版, 2004.
- [90] 山本圭治郎, 兵頭和人, 石井峰雄, 松尾崇. 介護用パワーアシストスーツの開発: 機械力学, 計測, 自動制御. 日本機械学会論文集. C編, Vol. 67, No. 657, pp. 1499–1506, May 2001.
- [91] 橋野賢. 移乗介助ロボットの現状と課題. 日本ロボット学会誌, Vol. 11, No. 5, pp. 49–54, Jul 1993.
- [92] 大森耕介, 伊藤京子, 酒田信親, 大西智士, 西田正吾. 他者との関わり意識を促す Synchronized Heart Beat System. ヒューマンインタフェースシンポジウム 2010 論文集, pp. 621–628, Sep 2010.
- [93] 佐々木正人. 複雑系の科学と現代思想アフォーダンス, 1997.
- [94] 埴淵俊平, 伊藤京子, 西田正吾. 笑顔の視覚化が会話参加者に与える影響 (HCS+フォーラム顔学ジョイントセッション, 顔とコミュニケーション). 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, Vol. 110, No. 247, pp. 23–28, Oct 2010.
- [95] 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.
- [96] 前川峻志, 田中孝太郎. Built with Processing: デザイン/アートのためのプログラミング入門. ビー・エヌ・エヌ新社, 2007.
- [97] 板倉昭二. 自己の起源: 比較認知科学からのアプローチ. 金子書房, 2012.
- [98] 松村真宏. 仕掛学概論: 人々の人々による人々のための仕掛学 (特集—仕掛学). 人工知能学会誌, Vol. 28, No. 4, pp. 584–589, Jul 2013.

- 
- [99] 神田崇行. HRIにおけるソーシャルロボット研究の動向. 日本ロボット学会誌, Vol. 29, No. 1, pp. 2-5, Jan 2011.
- [100] 佐々木正人. アフォーダンス—新しい認知の理論 (岩波科学ライブラリー 12) , 1994.
- [101] 山田誠二. 人とロボットの<間>をデザインする. 東京電機大学出版局, 2007.
- [102] 山田誠二, 寺田和憲, 小林一樹. 人を動かす HAI デザインの認知的アプローチ (特集—人を動かす HAI) . 人工知能学会誌, Vol. 28, No. 2, pp. 256-263, Mar 2013.
- [103] 吉田善紀, 吉池佑太, 岡田美智男. Sociable Trash Box: 子どもたちと一緒にゴミを拾い集めるロボット. ヒューマンインタフェース学会論文誌, Vol. 11, No. 1, pp. 27-36, Feb 2009.
- [104] 東村知子. あなたへの社会構成主義. ナカニシヤ出版, 2004.
- [105] 上野直樹. 協同的な活動を組織化するリソース. 認知科学—Cognitive Studies : Bulletin of the Japanese Cognitive Science Society, Vol. 3, No. 2, pp. 5-24, May 1996.
- [106] 大島直樹, 岡澤航平, 本田裕昭, 岡田美智男. TableTalkPlus: 参加者の共同性や社会的なつながりを引き出すアーティファクトとその効果. ヒューマンインタフェース学会論文誌, Vol. 11, No. 1, pp. 105-114, Feb 2009.
- [107] 大島直樹, 山口雄大, デシルバラビンドラ, 岡田美智男. Sociable Spotlight: コミュニケーションの「場」に対する構成的理解に向けて (「コミュニケーションと気づき」及び一般). 電子情報通信学会技術研究報告. HCS, ヒューマンコミュニケーション基礎, Vol. 111, No. 190, pp. 33-38, Aug 2011.
- [108] 三嶋博之. エコロジカル・マインド: 知性と環境をつなぐ心理学. 日本放送出版協会, 2000.

- [109] 大澤博隆. ヒューマンエージェントインタラクションの研究動向 (特集—エージェント). 人工知能学会誌, Vol. 28, No. 3, pp. 405–411, May 2013.
- [110] 岡田美智男. 口ごもるコンピュータ. 共立出版, 1995.
- [111] 岡田美智男. Talking Eyes—対話する「身体」を創る. システム/制御/情報: システム制御情報学会誌—Systems, Control and Information, Vol. 41, No. 8, pp. 323–328, Aug 1997.
- [112] 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, 2000.
- [113] 岡田美智男. コミュニケーションに埋め込まれた身体性—ロボット研究からのアプローチ (特集 コミュニケーションの身体性—からだは何を伝えるか). 言語, Vol. 37, No. 6, pp. 56–63, Jun 2008.
- [114] 岡田美智男. 人とロボットとの相互行為とコミュニケーションにおける身体性 (総特集メルロ=ポンティ—身体論の深化と拡張). 現代思想, Vol. 36, No. 16, pp. 300–311, Dec 2008.
- [115] 岡田美智男. ソーシャルなロボティクスと会話分析研究との接点を探る: 社会的相互行為に立ち会う視点の位置を巡って. 認知科学—Cognitive Studies: Bulletin of the Japanese Cognitive Science Society, Vol. 16, No. 4, pp. 487–493, Dec 2009.
- [116] 岡田美智男. 弱いロボット. 医学書院, 2012.
- [117] 岡田美智男, 片井修, 塩瀬隆之, 大須賀美恵子, 椎尾一郎. コミュニケーションと身体性. ヒューマンインタフェース学会論文誌, Vol. 5, No. 2, pp. 153–162, May 2003.
- [118] 志和敏之, 奥野佑将, 神田崇行, 今井倫太, 石黒浩, 萩田紀博. コミュニケーションロボットによる道案内: ジェスチャの有用性と発話タイミングのモデ

- ル化(人工知能, データマイニング). 電子情報通信学会論文誌. D, 情報・システム, Vol. 95, No. 10, pp. 1818–1828, Oct 2012.
- [119] 吉池佑太, 岡田美智男. ソーシャルな存在とは何か: Sociable PC に対する同型性の帰属傾向について (ヒューマンモデル, 特集—人とエージェントのインタラクション論文). 電子情報通信学会論文誌. A, 基礎・境界, Vol. 92, No. 11, pp. 743–751, Nov 2009.
- [120] 中野有紀子. 人とロボットの<間>をデザインする—第4章エージェントによるしぐさと視線のコミュニケーション. 東京電機大学出版局, 2007.
- [121] 山口雄大, 大島直樹, デシルバラビンドラ, 岡田美智男. ネットワークを介した促進行為場の生成について. ヒューマンインタフェース学会論文誌—The Transactions of Human Interface Society, Vol. 14, No. 1, pp. 415–424, 2012.
- [122] 星洋輔, 小林貴訓, 久野義徳, 岡田真依, 山崎敬一, 山崎晶子. 観客を話に引き込むミュージアムガイドロボット: 言葉と身体的行動の連携 (エージェントデザイン, 特集—人とエージェントのインタラクション論文). 電子情報通信学会論文誌. A, 基礎・境界, Vol. 92, No. 11, pp. 764–772, Nov 2009.
- [123] 米田隆志, 舟久保熙康. 車椅子・歩行介助装置の現状と課題. 日本ロボット学会誌, Vol. 11, No. 5, pp. 44–48, Jul 1993.
- [124] 寺田和憲. 人とロボットの<間>をデザインする—第8章 意図的な機械. 東京電機大学出版局, 2007.



## 謝辞

本研究の遂行ならびに論文の作成にあたり、懇切なる御指導を賜りました豊橋技術科学大学 情報・知能工学系 教授 岡田美智男 先生に謹んで感謝の意を表す。並びに、有益な御助言と御教示を賜りました豊橋技術科学大学 情報・知能工学系 助教 デシルバラビンドラ 先生に心より謝意を申し上げる。博士論文審査委員長である豊橋技術科学大学 情報・知能工学系 教授 堀川順生 先生、並びに審査委員である同系 教授 栗山繁 先生、准教授 金澤靖 先生には、ご多忙でありながら、審査を通して貴重なご意見、重要な指摘を頂いた。

本研究の遂行ならびに論文の作成にあたり、御協力いただいた、豊橋技術科学大学 情報・知能工学系 岡田研究室メンバの皆様感謝する。とくに、本田氏、竹井氏、吉池氏、吉田(善)氏、岡澤氏、鴨田氏、山地氏、福井氏、角氏、吉田(幸)氏、三宅氏、山口氏、森氏、巽氏、竹田氏、ユウセフ氏、吉田(広)氏、小田原氏、浅野氏、澤田氏、田中氏、堀田氏、上原氏、蔵田氏、深町氏、そして、依田和子氏には心から感謝している。次世代ロボット創出プロジェクト(後のTUTオープンチャレンジプロジェクト)でお世話になった藤井秀雪先生をはじめとする京都造形芸術大学のプロジェクトメンバーの皆様にも、感謝の意を表す。豊橋技術科学大学 電子情報工学専攻の上野氏、渋谷氏、中島氏とは同じ博士後期課程に在籍し、一緒に励ましあった。

筆者の研究するきっかけを作っていただいた国立長野工業高等専門学校 鈴木宏先生、伊藤祥一先生、堀内征治先生には、お会いする度に初心を思いだし、自分の原点に立ち戻ることができる。また、長野工業高等専門学校からの友人であり、豊橋技術科学大学における学生生活を共に過ごした、月岡氏、篠原氏にはと

でも元気づけられた。

最後に、筆者のことを誰よりも長く支えてくれた父の大島信，祖父の大島賢吾，祖母の大島キヨ，伯母の大島弘子，五十嵐孝子，叔母の中山福子に感謝したい。

2014年6月 大島直樹



# 本論文に関する研究業績

## 査読付学術論文

- [1] **大島直樹**, 山口雄大, デシルバ ラビンドラ, 岡田美智男: Sociable Spotlight : 社会的な関わりの中で構成されるアーティファクトについて, 人工知能学会論文誌, Vol. 29, No. 3-B, pp. 288–300 (2014).
- [2] 山口雄大, **大島直樹**, デシルバ ラビンドラ, 岡田美智男: ネットワークを介した促進行為場の生成について, ヒューマンインタフェース学会論文誌, Vol. 14, No. 4, pp. 415–424 (2012).
- [3] **大島直樹**, 岡澤航平, 本田裕昭, 岡田美智男: TableTalkPlus : 参加者の共同性や社会的なつながりを引き出すアーティファクトとその効果, ヒューマンインタフェース学会論文誌, Vol. 11, No. 1, pp. 105–114 (2009).

## 査読付レター

- [1] **Naoki Ohshima**, Yuta Yamaguchi, P. Ravindra S. De Silva, Michio Okada: Sociable Spotlights: The Symbolic Representation of Interconnectivity for Turn-Taking, *ICIC Express Letters*, Volume 6, Number 12, December 2012, pp. 2981–2988 (2012).

## 国際会議・国内学会・シンポジウム等における発表

- [1] 小田原雄紀, 蔵田洋平, 松下仁美, **大島直樹**, P. Ravindra S. De Silva, 岡田美智男: Talking-Ally: 聞き手性と宛名性に配慮した発話生成システムについて, *Human-Agent Interaction シンポジウム 2013 (HAI-2013) 論文集*, I-3, pp. 19–24 (2013).
- [2] 蔵田洋平, 小田原雄紀, 松下仁美, **大島直樹**, P. Ravindra S. De Silva, 岡田美智男: Talking-Ally における宛名性の表示機構について, *Human-Agent Interaction シンポジウム 2013 (HAI-2013) 論文集*, P8, pp. 77–80 (2013).
- [3] 蔵田洋平, 小田原雄紀, 松下仁美, **大島直樹**, デシルバラビンドラ, 岡田美智男: Talking-Ally: 聞き手性と宛名性に配慮した発話生成システム実現にむけて, *ヒューマンインタフェースシンポジウム 2013 DVD-ROM 論文集*, pp. 511–514 (2013).
- [4] Yuki Odahara, **Naoki Ohshima**, P. Ravindra S. De Silva, and Michio Okada: Talking Ally: Toward Persuasive Communication in Everyday Life, *Proc. of The 15th International Conference on Human-Computer Interaction (HCI 2013)*, pp. 394–403, Nevada, USA, 21–26 July (2013).
- [5] Yuki Odahara, Yohei Kurata, **Naoki Ohshima**, P. Ravindra S. De Silva and Michio Okada: Talking-Ally: Towards Persuasive Communication, *Proceedings of The 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2013)*, p. 413, March 3–6, Tokyo (2013).

- 
- [6] 小田原雄紀, 蔵田洋平, **大島直樹**, P.Ravindra S. De Silva, 岡田美智男: Talking-Ally: 聞き手性をリソースとする発話生成系の実現にむけて, *Human-Agent Interaction シンポジウム 2012 (HAI-2012) 論文集*, 2E-2 (2012). **HAI-2012 Outstanding Research Award 受賞**
- [7] **Naoki Ohshima**, Yusuke Ohyama, Yuki Odahara, P. Ravindra S. De Silva and Michio Okada: Talking-Ally: Intended Persuasiveness by the Utilizing Hearership and Addressivity, *Proc. of The fourth International Conference on Social Robotics (ICSR 2012)*, pp. 317–326, October 29–31, Chengdu, China (2012).
- [8] 蔵田洋平, 小田原雄紀, **大島直樹**, デシルバラビンドラ, 岡田美智男: Talking-Ally:聞き手と一緒に発話を組織する発話生成系の研究, 平成 24 年度電気関係学会東海支部連合大会講演論文集, D1-6 (2012).
- [9] 小田原雄紀, 蔵田洋平, **大島直樹**, デシルバラビンドラ, 岡田美智男: Talking-Ally : 聞き手と一緒に発話を組織する発話生成システムについて, *ヒューマンインタフェースシンポジウム 2012 論文集*, 1411L, pp. 161–166 (2012).
- [10] **Naoki Ohshima**, Yuta Yamaguchi, P. Ravindra S. De Silva and Michio Okada: Sociable Spotlights: Cognitive Artifacts to Enhance Engagement in Conversation, *Proc. of The 10th Asia Pacific Conference on Computer Human Interaction 2012 (APCHI 2012)*, pp. 411–416, August 28–31, Matsue, Japan (2012).
- [11] **Naoki Ohshima**, Yuta Yamaguchi, P. Ravindra S. De Silva, and Michio Okada: Sociable Spotlights: Cognitive Artifacts to Enhance Engagement in Conversation, *Proc. of The third International Conference on Social Robotics*

- 2011 (*ICSR 2011*), pp. 73–76, November 24–25, Netherlands (2011).
- [12] Naoki Ohshima, Yuta Yamaguchi, P. Ravindra S. De Silva, and Michio Okada: Sociable Spotlights: Cognitive Artifacts to Enhance Engagement in Conversation, *Proc. of The Asia-Pacific Interdisciplinary Research Conference 2011 (AP-IRC 2011)* p. 73, November 17–18, 2011, Toyohashi (2011).
- [13] 大島直樹, 山口雄大, デシルバラビンドラ, 岡田美智男: Sociable Spotlight: コミュニケーションの「場」に対する構成的理解に向けて, ヒューマンコミュニケーション基礎研究会 (HCS) & ヴァーバル・ノンヴァーバルコミュニケーション研究会 (VNV) 合同研究会, 信学技法, HCS2011-33, Vol. 111, No. 190, pp. 33–38 (2011).
- [14] Naoki Ohshima, Yuta Yamaguchi, P. Ravindra S De Silva and Michio Okada: Sociable Spotlights: A Flock of Interactive Artifacts, *Proceedings of 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2011)*, pp. 321–322, March 6–9, Switzerland (2011).
- [15] 山口雄大, 大島直樹, 巽孝介, P. Ravindra De Silva, 岡田美智男: NEXUS+: ネットワークを介した仮想クリーチャの共有による他者とのつながり感の創出, インタラクション2011 予稿集 (in CD-ROM) , pp. 331–332 (2011).
- [16] 大島直樹, 山口雄大, P. Ravindra De Silva, 岡田美智男: Sociable Spotlight: スポットライトはエージェントをもち得るか—傍観者としての参与の可能性, HAI シンポジウム 2010 資料集 (in USB memory), 2B-1 (2010).
- [17] 大島直樹, 山口雄大, P. Ravindra De Silva, 岡田美智男: 社会的相互行為を媒介するアーティファクトについて, 日本生態心理学会第3回大会プログラム

- (in CD-ROM), オープンフォーラム D-2 (2010).
- [18] **大島直樹**, 岡澤航平, 岡田美智男: Sociable Spotlight : 会話の場に参入する気ままなクリーチャ, インタラクシオン 2010 予稿集 (in CD-ROM), インタラクティブ発表・スタンダード SB30 (2010).
- [19] **大島直樹**, 岡澤航平, 岡田美智男: TableTalkPlus: 多人数会話を媒介するインタラクティブコンテンツとその会話に与える影響について, 電子情報通信学会ヒューマンコミュニケーショングループ (HCG) ヒューマンコミュニケーション基礎研究会 (HCS) HCS2009-49, Vol. 109, No. 224, pp. 5–10 (2009).
- [20] 岡澤航平, **大島直樹**, 宇都裕紀, 岡田美智男: TableTalkPlus : 参加者の共同性を引き出すインタラクティブコンテンツとそのデザイン, ヒューマンインタフェースシンポジウム 2009 論文集 (in CD-ROM), pp. 291–294 (2009).
- [21] 岡澤航平, **大島直樹**, 角裕輝, 吉池佑太, 岡田美智男: 会話のダイナミクスとその視覚化について, 日本生態心理学会第二回大会 発表論文集, pp. 23–24 (2008).
- [22] 岡澤航平, **大島直樹**, 角裕輝, 岡田美智男: VelicitA : 参加者の共同性を引き出すインタラクシオンデザイン, ヒューマンインタフェースシンポジウム 2008 論文集 (in CD-ROM), pp. 793–798 (2008).
- [23] **大島直樹**, 岡澤航平, 吉池佑太, 岡田美智男: 人と人との関わりを引きだすソーシャルメディアのデザインとその応用, Assistive Technology & Augmentative Communication Conference 2008, pp. 68–69 (2008).
- [24] 岡澤航平, **大島直樹**, 岡田美智男: TableTalkPlus による会話の「場」のデザ

イン, インタラクション 2009 論文集, 情報処理学会シンポジウムシリーズ, Vol. 2009, No. 4, D14 (2009).