

# 「時間」と「プロジェクション」: 時間的重なりによって創発する他者モデル “Time” and “Projection”: Models of Others Emerging by Temporal Overlapping

小野 哲雄<sup>\*1</sup>  
Tetsuo Ono

水丸 和樹<sup>\*1</sup>  
Kazuki Mizumaru

<sup>\*1</sup> 北海道大学 大学院情報科学研究科 ヒューマンコンピュータインタラクション研究室  
Human-Computer Interaction Lab., Graduate School of Information Science and Technology, Hokkaido University

To make it behave as if an artifact (agent) has an internal model (model of others) is important as a design principle for making people understand the artifact (AI system) with increased autonomy. In particular, in this research, we focused on "time" and proposed a model that allows people to assume other models of artifacts by temporal overlapping of utterance behaviors in communication. In this paper, first of all, after overviewing our research on "duality of the body" and "temporal overlapping" in the interaction, we proposed a framework of "models of others emerging by temporal overlapping". We conducted experiments based on this framework and verified its validity. In the future, from this point of view, we are planning to conduct research on the interaction design of people and artifacts based on "time".

## 1. はじめに

「プロジェクション科学」とは、人間の認知機構を解明するためのまったく新しい方法論であるとともに、そのモデル論的理解をとおして工学的応用までも射程に収めている野心的な構想である[小野 16]. [鈴木 16]が述べているように、認知科学におけるこれまでの研究は、刺激の「受容」とその利用過程に注目し、知覚、記憶、思考などのさまざまな認知現象について、精度の高いモデルを作り上げてきた。しかし一方で、内部に構成されたモデル(表象)が世界のどこかに「投射」(プロジェクション)されていると考えなければ理解できない現象も多数存在する。たとえば、視覚・聴覚などの(直接的な神経回路が存在しない)遠感覚における投射、ラバーハンドイリュージョン、子どもの遊び行動における投射(imaginary companion)、異常事態における他者の存在感(third man factor)などは新しい概念や機構を想定しなければ説明できないと考えられる。

本研究では、人とロボットのインタラクションにおける「時間」に注目し、「プロジェクション」との関係性を明らかにする。具体的には、2体のロボット同士のインタラクションにおいて、両者の発話行動を徐々に時間的に重ね合わせ(temporal overlapping)、予期的に振る舞わせる。これを観察していた人は、ロボットが相互に他者モデルを持っていると想定するようになり、両者がインタラクションを行っている空間を「社会空間」(social space)と認識するようになると予想している。このプロセスは後に述べるように、人間が「身体の二重性」(duality of embodiment)を持つがゆえに、自身の認知的なモデルをロボットへ「プロジェクション」することによって成立していると考えられる。

ここで本研究における「身体の二重性」とは、「外に向かう身体」と「内に向かう身体」を意味している。たとえば、人が他の人に道案内をする場合、外部へ表出するジェスチャという行為と自身の記憶にアクセスするプロセスは不可分であり、双方が影響を与え合っていることが明らかとなっている[Ono 00, 小野 01]。つまり、人は自身の「身体の二重性」を、予期的な振る舞いをしてるロボットにも付与するのではないかと予想している。

本研究の成果は、AI技術の進展により、自律性が高まった人工物(AIシステム)の挙動を人に理解させるためのインタラクションデザインの原理として重要である。つまり、人同士の信頼感の醸成は、お互いに外部に表出される情報だけから内部状態を推定し、その状態推定が「時間」的に抗えない相互の反応を繰り返すことによって行われる。言い換えると、同様の認知機構を持っているならば欺くことができない「時間的な制約」によってインタラクションを継続することによって信頼感は醸成されるのである。人とロボットのインタラクションにおいても同様に「時間」に基づくインタラクションデザインが必要であると考えられる。

## 2. 「身体の二重性」と「時間」

### 2.1 道案内における「身体の二重性」

本節ではまず、我々がすでに行った、人と人および人とロボットの道案内の場面におけるインタラクション実験における「身体の二重性」について述べる。[小野 01, Ono 01]では、身体同調的動作による対話者間の関係性の構築、および、この関係性に基づく情報伝達のメカニズムを理解するための枠組みを提案した。これらの研究における実験の結果、人と人のインタラクションでは、身体同調的動作により円滑な情報伝達が行われていることが明らかとなった(図 1 左)。また、人とロボットのインタラクションでは、人がロボットの身体の向きや、腕、首の動きに引き込まれるように同調的な動作をしたとき、円滑な情報伝達が行われることが明らかとなった(図 1 右)。一方、上記のような身体同調的動作が起こらない場合、人は道案内における空間情報の想起や記憶が難しいことが明らかとなった。

つまり、人が道案内で行うジェスチャ(「外に向かう身体」と空間情報の想起・記憶(「内に向かう身体」))は不可分であり、双方が「時間」的に同期しない場合、インタラクションが阻害されることが明らかとなった。したがって、これらの研究により、インタラクションにおける「身体の二重性」と「時間」の重要性が明らかとなった。



図 1. 道案内実験における人と人の同調的動作(左), 同様に人とロボットの同調的動作(右)

## 2.2 「時間的な重なり」とコミュニケーション空間

人とロボットのインタラクションにおいて、ロボットを文脈適応的に振る舞わせるためには、人同士のコミュニケーション空間やその活性度を認識することは重要である[小野 13]。この機能を実装することにより、ロボットが人同士のコミュニケーション空間を回避したり、対話の活性度が低下したタイミングで介入するなど、「空気を読む」ことにより円滑な社会的なコミュニケーションを実現するための基盤を構築することができる。

この機能を実現するために、我々是对話の参加者の音声発話の「重なり」に注目した対話活性度の計算方法を提案した[Imayoshi 12, Imayoshi 13]。この簡易な計算方法により、人同士のコミュニケーション空間を高い確率で認識できるようになり(図 2)、さらにこの計算方法を用いて、ロボットが人同士のコミュニケーション空間を認識し回避する応用システム(図 3)や、このコミュニケーション空間内だけで、情報漏洩することなく安全に人同士が対話することができるシステムを開発してきた。

これらの研究をとおして、音声対話における「時間的な重なり」はコミュニケーション空間を認識する重要な手がかりとなることが示された。

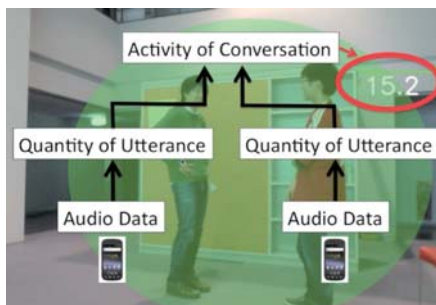


図 2. 音声発話の重なりを用いた対話活性度の計算



図 3. 社会空間を認識し、それを回避するロボット

## 3. 時間的な重なりによって創発する他者モデル

すでに述べたように、エージェントやロボットなどの人工物を内部モデル(他者モデル)を持っているかのように振る舞わせることは、自律性が高まった人工物(AI システム)の行動を人に理

解させるためのデザイン原理として重要である。本稿では、これまでの我々の研究(2 節参照)に基づき、発話行動の時間的な重なりによって人に他者モデルを想定させることができるインタラクションデザインの枠組み **Temporal Overlapping** を提案する。

図 4 に、本稿で提案する枠組みの概要を示す。2.1 節で述べたように、空間情報の円滑な伝達のためには、同調的な身体動作が重要な役割を果たすことが明らかとなった。また、2.2 節で述べたように、音声発話の時間的な重なりは、コミュニケーション空間を識別するための重要な指標となることが明らかとなった。つまり、発話と身体動作の「時間的な重なり」は円滑なコミュニケーションにおいては頻りに観察される現象であると考えられる。

2 体の人工物(ロボット)が上記のように「時間的な重なり」をもって発話と身体動作を行うとき、我々には 2 体が预期的に振る舞っているように見えるだろう。なぜならば、一方の発話や身体動作が終了する前に他方は行動を起こさねばならず、そこには「予期」(prediction)が不可欠となる。この予期的動作の基盤にあるのが、他者に対するモデル(他者モデル)であろう。この他者モデルは、同様の身体を有した者同士が、一定のインタラクションの履歴をとおして形成されていくものであると考えられる。

ここで、なぜ上記のような推論が成り立つのであろうか。それは我々が「身体の二重性」を有しており、実際に観察可能なのは「外に向かう身体」だけであるが、当然、不可分な「内へ向かう身体」があることを想定してしまっただけである。このプロセスは、自分の内部モデルを他者や人工物に無意識に「投射」(プロジェクション)することによって成立していると思われる。



図 4. インタラクションデザインの枠組みである **Temporal Overlapping** の概要

## 4. 検証実験

本研究で提案した枠組み **Temporal Overlapping** を検証するため、2 つの検証実験を行なった[水丸 17]。実験 1 では、2 体のロボットの発話行動の「時間の重なり」が人間に与える影響を調査するため、クラウドソーシングサービスを利用した。さらに、実験 2 では、ロボット 2 体のインタラクションを実際に実験参加者が観察を行い、発話行動の「時間の重なり」の効果を調査した。実験の詳細については、[水丸 17]を参照されたい。

### 4.1 実験 1 の実験設定と手続き

実験 1 では、ロボット(Pepper)2 体の発話の間隔が、それを観察する人間にどのような印象を与えるのかを調査するために、クラウドソーシングサービスを利用した。調査に使用したロボットのビデオは、発話間隔をパラメータとして、3 つの条件(Overlap 条件, Default 条件, Delay 条件)により作成された(図 5)。

まず、Overlap 条件では双方のロボットの発話が徐々に重なっていき、最終的には 1 秒重なるように設定した。会話ビデオ全体の長さは 60 秒であった。次に、Default 条件では発話間隔が常に 0 秒であり、一方のロボットの発話が終了するとすぐにもう

一方のロボットが発話を開始した。会話ビデオ全体の長さは 67 秒であった。最後に, Delay 条件では発話間隔は常に 1 秒であった。会話ビデオ全体の長さは 79 秒であった。

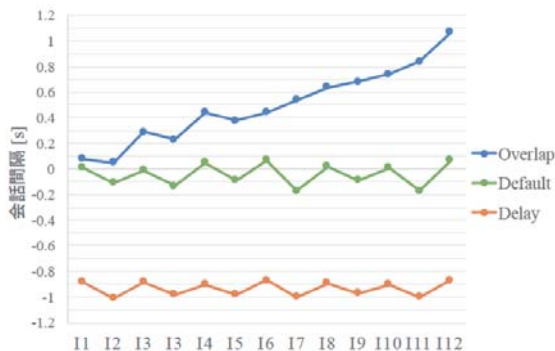


図 5. 各条件におけるロボットの発話間隔

実験 1 の実験参加者は 721 名であり, ロボット 2 体の会話ビデオを見た後, 以下の 3 項目のアンケートに 5 段階形式 (5 が最もポジティブ) に回答してもらった。

1. How natural was the conversation of the robots?  
(ロボットの会話はどの程度自然だったか)
2. How lively was the conversation of the robots?  
(ロボットの会話はどの程度活発だったか)
3. How familiar were the robots with each other?  
(ロボット 2 体の関係はどの程度親密だったか)

#### 4.2 実験 1 の結果

実験 1 のアンケート結果について, 1 要因被験者間比較 (クラスカル・ウォリス検定) を行なったところ, 項目 2「ロボットの会話はどの程度活発だったか」について有意傾向が確認された (図 6)。このため, 事後検定として Mann-Whitney の U 検定 (片側検定) を行なったところ, Overlap 条件と Delay 条件との間に有意な差が確認された ( $p < .05$ )。この結果は, 発話の間隔が常に 1 秒ある Delay 条件よりも, 発話の重なりが生じる Overlap 条件の会話の方が活発に感じることを示している。

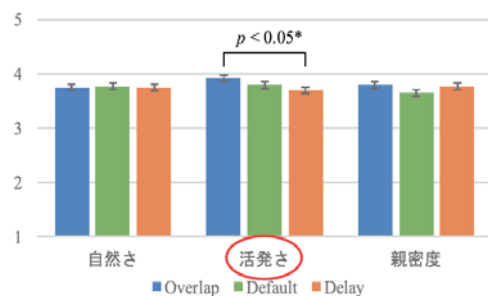


図 6. 実験 1 のアンケート結果

#### 4.3 実験 2 の実験設定と手続き

実験 2 では, 実際のロボット 2 体を用いて, ロボット同士の会話を実験参加者が直接観察した (図 7)。また, 実験手続きは以下のとおりである。

- 1) 実験参加者は実験の説明を受け, 2 体のロボットの前に立つように指示を受ける
- 2) 実験者の合図で実験参加者はロボットに挨拶する
- 3) ロボットの会話が始まる
- 4) 実験者は別室で様子をモニタリングする

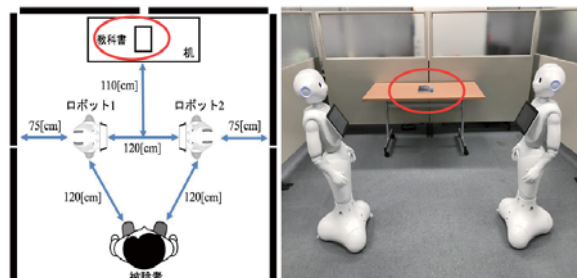


図 7. 実験 2 の実験環境 (左: 平面図, 右: 実験参加者視点)

5) 会話の終わりに片方のロボットが実験参加者に机の上の教科書を見るようにすすめる

6) 実験者は会話終了後, およそ 30 秒後に実験室に入り, 実験終了の旨を伝える

7) 実験参加者はアンケートに答える

手順 7 において, 実験参加者は以下の 5 段階形式 (5 が最もポジティブ) のアンケートに答えた。

1. ロボットの会話はどの程度自然だったか
2. ロボットの会話はどの程度活発だったか
3. ロボット 2 体の関係はどの程度親密だったか

#### 4.4 実験 2 の結果

実験 2 のアンケート結果については, どの項目間についても有意な差は確認されなかった。ところが実験参加者の行動を分析したところ, 行動は以下の 2 つのパターンに分けられた。

1. 教科書を見に行くときに, ロボット 2 体の間を通る
2. 教科書を見に行くときに, ロボット 2 体の間を避け, いずれかのロボットの後ろを通る

条件ごとに実験参加者がどちらの行動をとったか人数を計測した (表 1)。フィッシャーの正確確率検定により有意な差が確認されたため, 事後検定を行なった。その結果, Overlap 条件と Delay 条件との間に有意な差が確認された ( $p < .05$ )。この結果は, Delay 条件よりも, 発話の重なりが生じる Overlap 条件の方がロボットの間を避けて通る実験参加者が多いことを示している。

	ロボットの間を通る	ロボットの間を避ける
Overlap	4	9
Default	8	5
Delay	11	2

表 1. 各条件における実験参加者の行動人数

#### 4.5 考察

本節では本検証実験における実験 1 と実験 2 の実験方法の違いによる影響, および Overlap 条件と Delay 条件の結果の違いについて考察を行う。

まず実験 1 と実験 2 は本質的に異なる実験であると考えられる。実験 1 はロボット 2 体の発話行動の「時間の重なり」が人間に与える影響を調査することを目的としていた。実験結果から当初の実験の目的は達成されているが, ロボットが本来持つ身体性 (インタラクションにおいて人間の身体を引き込むような動的な性質) は当然含まれてはおらず, より客観的な視点から印象評価がなされていたと思われる。一方, 実験 2 は上記の意味でのロボットの身体性を含む, 実験参加者の主観的な視点からロボットのインタラクションを捉えていたため, 表 1 で示したような

行動が生起したと思われる。このメカニズムを探ることは大変興味深いため、5 節において改めて議論したい。

## 5. 「時間」に基づくインタラクションデザイン

人工物(エージェント)が内部モデル(他者モデル)を持っているかのように振る舞わせることは、自律性が高まった人工物(AI システム)を人間に理解させるためのデザイン原理として重要である。特に本研究では「時間」に注目し、発話行動の時間的重なりによって人に人工物の他者モデルを想定させることができるインタラクションデザインの枠組み **Temporal Overlapping** を提案した。4 節で述べた実験結果に基づき、この提案した枠組みの意味について議論したい。

現時点での仮説的な説明を図 8 に示す。この図を簡潔に説明すると以下のとおりとなる。

1. **Overlap** 条件の実験参加者は、発話行動が徐々に重なっていく 2 体のロボットを観察した(図中のロボットの Utterance 間にある **Overlapping** の部分の変化に注意)

2. 発話行動の時間的重なりから、2 体のロボットが予期的に振舞っているように実験参加者には見える(図中の **Prediction** を認識するようになる)

3. 上記の 2 より、ロボットに「他者モデル」を想定するようになる

この理由としては、実験参加者は同調的な発話行動を行う 2 体のロボットを観察するだけであるが、人間には「身体性の二重性」(「外に向かう身体」と「内に向かう身体」)があるため、それを無意識にロボットに「投射」(プロジェクション)してしまい、ロボットにも「他者モデル」(「内に向かう身体」)があると想定してしまうからであると推測できる。

上記 1~3 のプロセスをとおして、実験参加者は 2 体のロボットに他者モデルを想定し、さらにロボット間に社会空間を認識したため、**Overlap** 条件ではロボット間を通るのを避け、ロボットの後ろを通ったのではないかと考えられる。

この社会空間の認識は、インタラクションに参加している人の主観的な時間(「いま」)においてのみ立ち現れると考えられる。現在、このプロセスのより詳細な検討を行っているが、「時間」と「プロジェクション」の関係は人間の認知における本質的なメカニズムであると思われる。今後はこのメカニズムの解明とともに、その工学的な利用を考えていきたい。

すでに述べたように、人工物が内部モデルを持っているかのように振る舞わせることは、自律性が高まった人工物(AI システム)を人に理解させるためのデザイン原理として重要である。今後は、この観点から、「時間」に基づく人と人工物のインタラクションデザインについて研究を進めていく予定である。

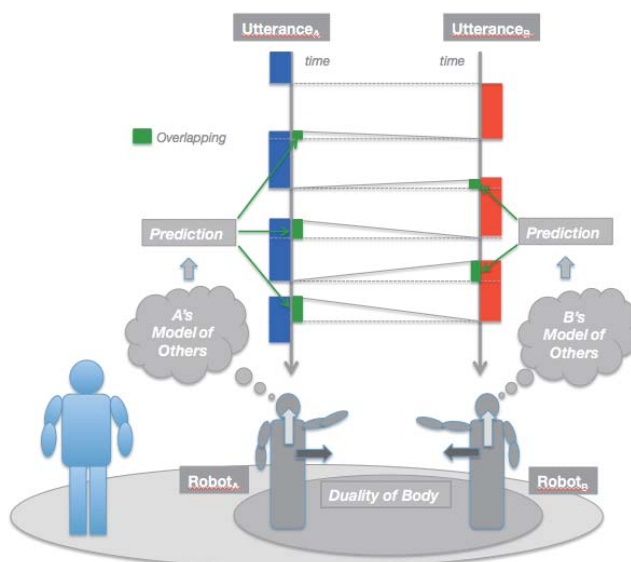


図 8. **Temporal Overlapping** と実験結果の仮説的説明

## 参考文献

- [Baker 17] Chris L. Baker, Julian Jara-Ettinger, Rebecca Saxe and Joshua B. Tenenbaum, Rational quantitative attribution of beliefs, desires and percepts in human mentalizing, *Nature Human Behaviour* 1, pp. 1-10, 2017. DOI: 10.1038/s41562-017-0064.
- [Imayoshi 12] Imayoshi, A., Munekata, N. and Ono, T.: Augmented Social Space: Robots Can Generate Context-Aware Adaptive Behaviors According to the Activity of Communications, *Proc. the 2012 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS 2012)*, Workshops and Tutorials, TW8\_0011, 2012.
- [Imayoshi 13] Akira Imayoshi, Hiroshi Yoshikawa, Nagisa Munekata, Tetsuo Ono. Robots that Can Feel the Mood: Adaptive Interrupts in Conversation Using the Activity of Communications, *Proc. of 1st International Conference on Human-Agent Interaction (iHAI 2013)*, 2013.
- [水丸 17] 水丸和樹, 坂本大介, 小野哲雄 (2017). 複数ロボットの発話の重なりによって創発する空間の知覚, HAI シンポジウム 2017, G-15, 2017.
- [小野 01] 小野哲雄, 今井倫太, 石黒浩, 中津良平. 身体表現を用いた人とロボットの共創対話, 情報処理学会論文誌, Vol. 42, No. 6, pp. 1348-1358, 2001.
- [Ono 01] Tetsuo Ono, Michita Imai, Hiroshi Ishiguro. A Model of Embodied Communications with Gestures between Humans and Robots, *Proc. of 23rd Annual Meeting of the Cognitive Science Society (CogSci2001)*, pp. 732-737, 2001.
- [小野 13] 小野哲雄, 今吉晃. 「空気を読むロボット」: コミュニケーション空間を利用した人を動かす HAI デザイン, 『人工知能学会誌』, Vol. 28, No. 2, pp. 284-289, 2013.
- [小野 16] 小野哲雄: 「プロジェクション・サイエンス」の視点からの認知的メカニズムのモデル論的理解, 2016 年度日本認知科学会第 33 回大会, OS03-2, 2016.
- [鈴木 16] 鈴木宏昭: プロジェクション科学の展望, 2016 年度日本認知科学会第 33 回大会, OS03-1, 2016.