

コミュニケーションロボットにおけるジェスチャ表出の有効性

Effectiveness of Hand Gestures in Communication Robots

福原 佑貴^{*1}
Yuki Fukuhara

上出 純平^{*2}
Junpei Kamiide

中野 有紀子^{*2}
Yukiko Nakano

^{*1} 成蹊大学 理工学専攻理工学研究科
^{*1} Graduate School of Science and Technology

^{*2} 成蹊大学 理工学部 情報科学科
^{*2} Dept. of Computer and Information Science Seikei University

This study examines whether robots showing appropriate gestures in proper timing are more successful in eliciting user's natural response, such as eye gaze and head nod, in human-robot communication. While many studies in robot gestures have focused on deictic gesture, we investigate how people perceive iconic gestures and baton gestures used as interactive gesture. We set up four experimental conditions: human, gesture synchronization robot, gesture a synchronization robot, and speech only robot. The results indicated that people pay more attention to iconic gestures which are executed in proper timing, but similar results were not found in baton gestures. Moreover, people slightly more frequently nodded to the robots in gesture synchronization condition, but this result was not statistically significant. Although the results found in this study are still weak, the trends of the results support our hypothesis.

1. はじめに

近年、ロボット技術の発展に伴い、特定の仕事のみを行う産業用ロボットだけでなく、人と直接関わり、共生することを目的としたコミュニケーションロボットの研究が注目されている。コミュニケーションロボットの多くは、人と同様の身体性を持つヒューマノイド型ロボット(図1)によるものであるが、その理由として、人はインタラクションする対象に対して擬人化を行う傾向が強いことがあげられる [Reeves & Nass 96]。つまり、人との関わり合いを目的としたロボットにおいて、この特性を最大限に利用することにより、人とロボットとの自然なコミュニケーションを実現しようとしている。また、そのためには、身体的な特徴が類似しているだけでなく、ロボットの振る舞いを人に近付けることが重要であると考えられる。

ロボットによる身体性を用いた表現の1つとして、ジェスチャがある。人は対話においてジェスチャを用いることにより、伝えるべき内容に関連のある情報を表し、会話の内容を補足している。従って、ロボットがジェスチャを効果的に使用できることは、人対ロボットとのコミュニケーションにおいても重要であると考えられる。既に、先行研究ではロボットによるジェスチャが人とのコミュニケーションにおいて有用であることが示されているが、それらはロボットによる指差し動作に関するものが多かった[Okuno 09]。しかし、指差しのほかにも、ものの形や大きさを表すジェスチャ等、人が用いるジェスチャの種類や機能は多様である。

一方、コミュニケーションロボットと同様、人間と類似した身体性を持つ人工物として、アニメーションエージェントがある。[1]ではアニメーションエージェントによるジェスチャの有効性について評価実験を行い、エージェントがジェスチャを行うことにより、被験者はそのエージェントをより協調的で、言葉も自然であるという印象を持つと報告



図 1: コミュニケーションロボット Nao

連絡先: 福原佑貴, 成蹊大学 理工学専攻理工学研究科
〒180-863, 東京都武蔵野市吉祥寺北町 3-3-1
TEL: 0422-37-2104, E-mail: dm106221@cc.seikei.ac.jp

している。そこで本研究では、Wizard-of-Oz によるロボットと人との対話実験を実施し、被験者の視線とうなずき動作を分析することにより、ロボットによる指差し以外の様々な形態のジェスチャが適切に用いられることが人対ロボットの自然なコミュニケーションを実現するために重要であることを検証する。

2. Wizard-of-Oz 実験環境

Wizard-of-Oz 実験を実施するために図1に示すヒューマノイド型ロボット(Aldebaran 社製 Nao)を用いた実験環境を構築した。Nao は、全長 58cm, 25 自由度の小型ロボットであり、ジェスチャの表現には問題ないと判断した。

2.1 音声とジェスチャの同期

ロボットの発話音声ファイルとジェスチャの動作スケジュールが記されたスクリプトファイルは CAST[Nakano, 04]を改良し、生成した。テキストが入力されるとそれを解析し、言語情報からジェスチャのタイミングを決定し、エージェントのアニメーションスケジュールを作成するとともに、合成音声を用いて発話音声を生産する。

Wizard-of-Oz 実験では GUI によりロボットの発話を指定すると、ロボットに音声ファイルが送信され、再生される。これと同時に、スクリプトファイルに記述された動作スケジュールに従って、ロボットに動作命令が送信され、指定されたタイミングでロボットのジェスチャが実行をおこなえるようにした。図 2 にロボットの発話と動作管理の構成を示す。

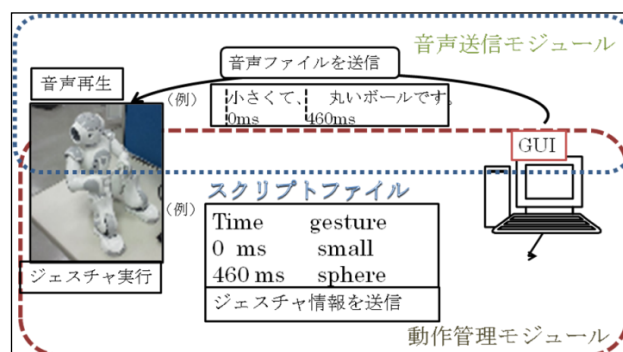


図2: ロボットの発話と動作管理

3. ロボットによるジェスチャ

本実験ではロボットに[McNeil 92, Bevalas 92]を参考に以下の2つのジェスチャを実装した。

(1) Baton ジェスチャ:発話の最後に対話の相手に手を差し伸べるような動きであり、相手に発話権を譲渡するジェスチャ。

(2) Iconic ジェスチャ:物体の形状、人や物の動作を表すことにより会話の内容を補足するジェスチャ。

これらのジェスチャをロボットがおこない、ジェスチャが意味のある形で人に伝われば、会話を円滑に進めることができ、コミュニケーションのパートナーとなるロボットの実装に役立つと考えられる。

4. 実験

4.1 実験の概要

被験者は人および、3つの条件で動作するロボットと対話を行う。ロボットが適切なタイミングでジェスチャをすることにより、被験者が人に対してと同様の反応をロボットに対して行うのかどうかを検証する。

本実験では、被験者の反応の指標として、ジェスチャへの注目比率とうなずき回数をを用いる。人同士のコミュニケーションにおいて、視線は会話への参加や理解を示すだけでなく、ユーザの興味・関心の対象をうかがい知る手がかりにもなる。また、うなずきは相手の発話を認識し、かつ会話の内容を理解したことを相手に伝えるジェスチャとして行われるものである。従って、うなずき行為は、被験者とロボットとのコミュニケーションの自然さを評価する指標になると考えた。

4.2 実験手順

被験者には、粘土で作成した動物の顔の説明をロボット(又は人)と交互に行うという設定で対話を行ってもらった。説明の対象となる顔部分は「鼻、眼、頬、口」の4種類である。セッション開始前に被験者に眼と口の形を見せ、それを覚えてもらった。従って、ロボット(または人)が鼻の形を説明すると、次に被験者が目の形を説明し、その後、ロボットが頬を、最後に被験者が口の形を説明すると言った流れで、交互に顔部分の説明をすることにより会話が進められる。被験者には、会話中に説明を受けた顔部分の形を覚えるよう指示し、被験者が対話の内容を理解しているかを確認するため、対話の最後に被験者からロボット(又は人)に記憶した4種類の顔部分の形を説明してもらった。



図 3: 実験の様子

4.3 実験条件

被験者は以下の4条件で対話を行った。

- (1) 人間条件: 実験者が被験者と対話を行う。ジェスチャは、実験者により適切なタイミングで使用されていると仮定する。
- (2) ジェスチャ同期条件: ロボットが発話中の適切なタイミングで適切なジェスチャを行う。
- (3) ジェスチャ非同期条件: ロボットが発話内容とは関係のないジェスチャを不適切なタイミングで行う。
- (4) 音声条件: ロボットは音声のみでジェスチャを行わない。

4.4 収集データ

実験にはロボットに関する予備知識のない大学生15名が参加した。尚、実験は各条件1回ずつとし計4回(各実験は4~5分程度)実施した。

視線情報の取得にはアイトラッカ(TobiiX120)を用いて、注視対象となる人やロボットに対する視線情報を取得し、ビデオカメラで被験者の顔映像を撮影した。また、同期条件と非同期条件の対話実験終了後には、被験者にロボットの印象に関するアンケートに答えてもらった。

5. 仮説

どの条件のロボットに対して、被験者はより強く擬人化を行い、自然なコミュニケーションが実現するのかを実験により検証するために、無意識に行う注視行動や傾きを被験者の反応指標とした。被験者が、人と同じタイミングでジェスチャを行うロボットに対して自然なコミュニケーションを行うのであれば、ジェスチャ同期条件において、人間に対してと最も類似した反応を返すであろう。もし、ロボットの動作タイミングに関わらずロボットが動くことに対して擬人化を行うなら、ジェスチャ同期条件とジェスチャ非同期条件間で被験者の行動に差はないであろう。また、ロボットの動作を煩わしいものと捉えるならば音声条件において、人に対してと類似した反応を返すと予想される。

6. 分析

本実験のデータ分析には、対話相手(ロボット/人)へのジェスチャ注視行動と、発話中とその後2秒以内のうなずきを分析データとして用いた。

6.1 ジェスチャの注視判定

本実験では、被験者がジェスチャを注視しているか否かの判定に Tobii Studio による AOI (Area of Interest) 分析ツールを用いた。AOI 分析ツールでは、事前に領域を定義しておき、領域内にアイトラッカにより取得された視線情報が入った場合に、その領域を注視していると判定する。この情報を用いて、説明者がジェスチャをおこなっている時間中に、被験者が注視している AOI がジェスチャの振る舞い領域内であれば、つまり、腕部の移動領域であれば、被験者は説明者がおこなっているジェスチャを注視していると定義した。図 4 に実際に実験で用いた AOI 領域の定義画面を示す。



図 4: 人とロボットでの AOI の定義画面

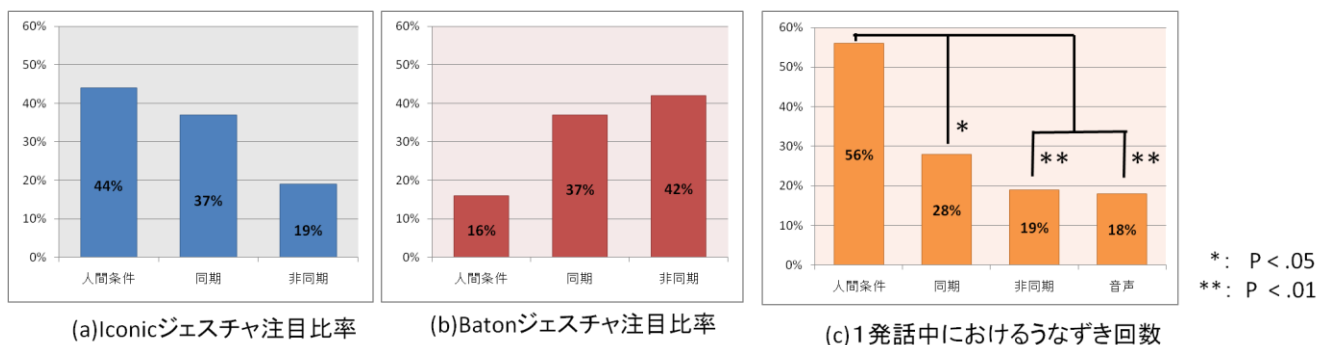


図 6：各ジェスチャの注目比率とうなずき回数

また、被験者の顔を撮影した映像から、対話者の発話中もしくは発話終了2秒後までに観測された被験者のうなずき回数を数えた。図5に実際に使用した分析の様子を示す。



図 5：視線情報とユーザのうなずき情報をアノテーションツールにて統合した様子

7. 結果

7.1 ジェスチャ注視の分析

全被験者のジェスチャの注目比率の平均値を図 6(a), (b)に示す。Iconic ジェスチャへの注目比率の平均値に対し、繰り返しありの一元配置分散分析を行ったところ、実験条件による有意差が見られた ($F(2, 42)=3.340386, p<.05$)。多重比較では、条件間の有意差は見られなかったが、ジェスチャ同期条件の Iconic ジェスチャ注目比率(37%)は人間条件(44%)に近い値を示した。

また、全被験者の Baton ジェスチャへの注目比率の平均値についても同様の分析を行ったところ、実験条件による有意差が見られた ($F(2, 42)=3.340386, p<.05$)。多重比較を行ったところ、それぞれの条件間での有意な差は見られなかった。また、人間条件の値が極端に低くなっていることがわかるが、このような結果となった理由として、Baton ジェスチャは Iconic ジェスチャとは異なり、ジェスチャ自体は意味を表現するものではないので、人同士の対話においては Baton ジェスチャにはあまり注目されなかったが、ロボットが行う Baton ジェスチャの動作が不自然に大きかったため、人間のジェスチャ以上に注目を集めてしまったためと考えられる。上記の考察から、ロボットの Baton ジェスチャはより小さい動きとして実装する方が望ましい結果が得られた可能性がある。

うなずきの分析

次に、対話相手の発話中とその後2秒以内における、被験者の平均うなずき回数を図 6(c)に示す。繰り返しありの一元配置分散分析の結果、実験条件による有意差が見られた ($F(3, 56)=2.76943, p<.05$)。多重比較を行ったところ、人間条件とその

他の条件間のみ有意な差が見られた。このことから、対話相手が人とロボットではうなずき動作を行う回数に差があることが分かる。また、ロボットの3条件間で比較すると、ジェスチャ同期条件で最もうなずき回数が多く、他の条件のロボットに対してよりも、より人に対するのに近い反応を示したと言える。

7.2 アンケートデータの分析

同期条件と非同期条件の対話実験終了後に被験者にアンケートに答えてもらった。アンケートの内容はジェスチャの有効性、ロボット全体の印象としての自然さ、ジェスチャの自然さの3項目であり、各項目について5段階でのリッカート法で評価するものである。各項目とアンケートの結果については図7に示す。各項目について Wilcoxon の符号付き順位検定を行ったところ、条件間でジェスチャの有効性において有意差が見られた ($p<.01$)。また、ジェスチャの自然さで有意差傾向が見られた ($p<.1$)。

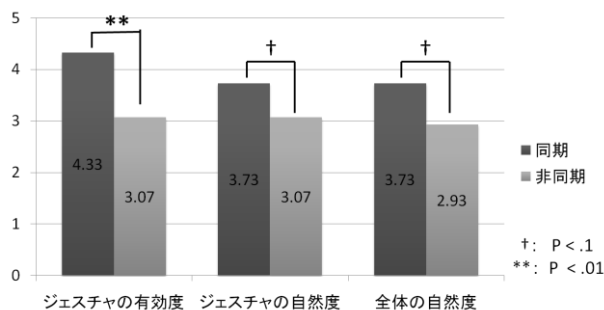


図7:アンケート結果

8. まとめ

本研究では、ヒューマノイドロボットが iconic ジェスチャやインタラクティブジェスチャ[Bavelas 92]としての baton ジェスチャを適切なタイミングで行うことにより、ユーザがロボットに対してより強い擬人化を行うようになり、ロボットが人間の自然な反応を引き出すことができるかどうかを、ロボットと人の対話実験を実施して検証した。その結果、Iconic ジェスチャの注目比率とうなずき回数に関して、ジェスチャ同期条件において最も人間に対する反応に近い結果が得られた。Baton ジェスチャに関しては、ジェスチャを必要以上に被験者の目を引く動きとなってしまったために、予想通りの結果とはならなかったが、Iconic ジェスチャの結果より、Baton ジェスチャについても同様の結果を得られる可能性があると考えている。

さらにアンケート結果からも、人は単にロボットが動くことに反

応しているのではなく、適切なタイミングで動作するロボットの動きをコミュニケーションシグナルとして認識し、肯定的な印象を持つことが示された。これらの結果は、ロボットが適切なタイミングでおこなうジェスチャは、ロボットが人のコミュニケーションのパートナーとなるために重要な要因であることを示唆している。

今後は、Baton ジェスチャと Iconic ジェスチャ以外のジェスチャに対しても、同様の結果が得られるのかを検証する必要がある。また、今回は視線情報についてジェスチャ区間内での AOI の分析のみをおこなったが、今後は視線の遷移パターンや視線の移動量等、行動指標を追加して、詳細な分析をおこなうことも重要である。

参考文献

- [Reeves 96] Reeves, B. and C. Nass, *The Media Equation: How People Treat Computers, Televisions and New Media Like Real People and Places*. 1996: CSLI Publications.
- [Okuno 09] Okuno, Y., et al., *Providing route directions: design of robot's utterance, gesture, and timing*, in the 4th ACM/IEEE international conference on Human robot interaction. 2009, ACM. p. 53-60.
- [Nakano 04] Nakano, Y.I., et al. *Converting Text into Agent Animations: Assigning Gestures to Text*. in *Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2004), Companion Volume*. 2004. Boston.
- [McNeill 92] McNeill, D., *Hand and Mind: What Gestures Reveal about Thought*. 1992, Chicago, IL/London, UK: University of Chicago Press.
- [Bavelas 92] Bavelas, J.B., N. Chovil, and e. al, *Interactive Gestures*. *Discourse Processes*, 1992. **15**: p. 469-489..