

ヒューマノイドロボットにおけるスピーカの位置の重要性

The importance of the location of loudspeaker for humanoid robot

金井 祐輔*1

Kanai Yusuke

今井 倫太*2

Imai Michita

*1慶應義塾大学大学院理工学研究科

Graduate School of Science and Technology Keio University

*2慶應義塾大学理工学部

Faculty of Science and Technology Keio University

Recent researches in robotics have developed various communication robots which can cooperate and communicate with humans. Various actuators and sensors have enhanced the research of robot's movement and appearance for human-robot interaction. However, there are little concerned about the location of a speaker in designing the robot. Humans speak with their mouth, while many of developed robots have a loudspeaker on their breast or abdominal region. This paper has conducted an experiment to confirm whether the location of a robot's loudspeaker has a effect on the human's understanding of an utterance. The experiment results have indicated that the robot's loudspeaker mounted on the head part improves the humans understandings of the utterance rather than the one abdominal part.

1. はじめに

ロボティクスの研究は、人の代わりに特定の作業を遂行するロボットの開発から人と直接かかわりを持ちコミュニケーションをとるロボットの開発へと発展した。現在までに開発された人とコミュニケーションするロボット（コミュニケーションロボット）には個人の生活を支援することを目的としているものがあり、例えば人を癒すことでメンタルケアをおこなったるロボット（支援ロボット）[7]、人を楽しませるロボット（エンターテインメントロボット）[7, 2]、また対話によって情報を提供するロボット（案内ロボット）[1]がある。本研究では、音声対話をするコミュニケーションロボットのデザインを音声発話の観点から考える。

人とロボットの音声対話に関する研究ではロボットの音声や言語表現だけでなく、人とロボットの実世界の情報の共有のための身体表現の研究も行われている。コミュニケーションロボットと人との間で実世界の情報を参照するための音声対話のデザインの研究は従来から行われている。例えば、コミュニケーションロボットの視線や適切な指差し、発話が人に実世界の情報を適切に誘導できることが明らかになっている[3, 10, 11]。従来研究の示すコミュニケーションロボットの視線や適切な指差し、発話の効果は人同士での対話においてジェスチャーや視線が重要であると同様に、人とコミュニケーションロボットの対話においてもロボットのジェスチャーや視線の重要性を示している。

コミュニケーションロボットの音声対話におけるロボットの発話や身体表現に関する研究はなされているのに対し、発話そのものが出力される音源の位置についての検討はあまり行われていない。現在開発されている人型ロボットの多くは胸部や腹部にスピーカを持ち、口から声を発する人とは異なる。

本研究ではロボットのスピーカの位置による人の発話理解への影響について検証する。従来開発された多くのロボットと同様に腹部のスピーカから発話する場合と、人と同様に頭部のスピーカから発話する場合の違いを比較検討する。

本稿の章立ては以下のとおりである。2章では人とロボットの実世界を参照した対話における視線、指差し、発話のデザインの重要性を説明したのち、ロボットのスピーカの位置に関する検討の必要性を指摘する。3章では上記で説明した実験について詳細な手順や被験者のデータについて述べる。また、ロボットのスピーカの位置が人の理解に影響を与えるということを実験で得られたアンケート結果から調査する。4章では考察を述べ、5章でまとめを述べる。

2. 実世界の情報を参照した人とロボットの対話

2.1 CG エージェントとの対話/ロボットとの対話

ロボットはCG エージェントとは異なり実世界に身体を持つ。実在する身体を用いてロボットは、人同士のコミュニケーションに近い形でより自然に実世界の情報を参照することが出来る。

実世界に関する対話において Shinozawa らは、実世界の情報を参照する場合はCG エージェントよりもロボットの方が有利であることを示した[6]。

2.2 実世界の情報の参照

ロボットと人との間で実世界の情報を参照するためのコミュニケーションの研究は従来から行われている。実世界の情報表現においてジェスチャーや視線といったロボットの身体表現が利用されている。人同士のコミュニケーションでジェスチャーや視線による非言語コミュニケーションは重要であり、[8]、その重要性は人とロボットのコミュニケーションにおいても同様であることが明らかになりつつある。例えば、ロボットの視線[3]や指差し、発話のデザイン[10, 11]の重要性が挙げられる。

ロボットの視線の重要性は、人の注意を実世界の情報に向ける際に顕著である。ロボットが注意を向けているものと同じものに人の注意を向けて欲しいというとき、ロボットと人がアイコンタクトを行うことが重要である。[3]。

今井らはロボットと人、物体の位置関係から、人の注意をその物体に引き込む視線の動きを生成する機構を開発し[10]、さらに杉山らはロボット・人・複数の物体の位置関係から指差しの方向と適切な指示語を選択・決定するシステムを開発した

連絡先: 金井祐輔, 慶應義塾大学大学院理工学研究科, 神奈川県横浜市港北区日吉3丁目14-1, TEL/FAX:042-388-1571, kana@ayu.ics.keio.ac.jp

[11]. ロボットが適切な指差しと指示語を用いることにより複数の物体が存在する中でロボットが物体への人の注意を誘導できる.

2.3 ヒューマノイドロボットのデザインの対話への影響

ヒューマノイドロボットのジェスチャや視線の動きをデザインすることで、人とロボットが実世界の情報を容易に共有できることが従来研究から分かる. しかし、ロボットとの音声対話では、発話は音声でなされるのにたいして、ロボットの音声の音源の位置に関する検討は皆無である. 本稿ではロボットの音声の音源の位置による人の実世界の情報の共有への影響を示す. 人と同様に頭部から発話音声が出力される場合と、多くの人型ロボットと同様に腹部から出力される場合を比較検討する. 特に、頭部にスピーカがある場合には、視線の移動に応じて音声の出力方向が変化するので、ロボットが視線を動かしながら対話する場面、例えばロボットが視線を動かすことによって実世界の情報に注意を向けていることを人に認識させる場合に強く影響がでると予想できる. そこで本稿では、ロボットがある物体に視線を向けながら、物体を参照する発話を行う場面に注目する. 頭部と腹部におけるロボットのスピーカの位置の違いの影響を、物体参照に対する人の注意対象を比較する事で明らかにする.

3. 実験

3.1 実験の目的と仮説

本稿ではロボットのスピーカの位置を頭部に設置する場合と腹部に設置する場合とで、ロボットが参照する物体を人が把握するのに差が生じるかを確かめる. ロボットは、視線(頭部)の動きと発話で物体の参照を行う. もし、視線の動きのみで人が正しく物体を参照することが可能ならば、音声の音源の位置に影響されずに被験者は参照の対象である物体に注意を向けるはずである. 一方で、音源の位置の影響があるのであれば、視線の動きのみでは正しく参照対象を伝えられないことになる.

本稿で検証される仮説は次のものである.

仮説 スピーカの位置を頭部に設置する場合と腹部に設置する場合でロボットからの物体参照に対する被験者の理解に差が生じ、頭部にある方が正しく理解されやすい.

3.2 実験環境

実験で用いるロボットは国際電気通信基礎技術研究所が開発されたコミュニケーションロボット「Robovie」[4]を用いている. 本研究では Robovie の頭部と胴体の内部にスピーカを設置されている. 頭部のスピーカは Robovie の口に相当する位置にあり、人から目視可能である. 腹部のスピーカは胴体内部にあるので目視できない.

実験は慶應義塾大学理工学部内の地下2階で行われた. また実験環境は図1に示すように Robovie の前に2つの箱を置き、被験者と Robovie が対面する. ここで Robovie から見て左側にある箱を箱1(図2の box1)、右側にある箱を箱2(図2の box2)とする.

3.3 実験条件

本研究では2つの実験条件を設定した.

実験条件1 頭部に設置したスピーカから Robovie が発話する.

実験条件2 腹部に設置したスピーカから Robovie が発話する.

2つの実験条件のもとでロボットの指示に対する被験者の行動を比較する. また実験条件の対照性を維持するため2つのス

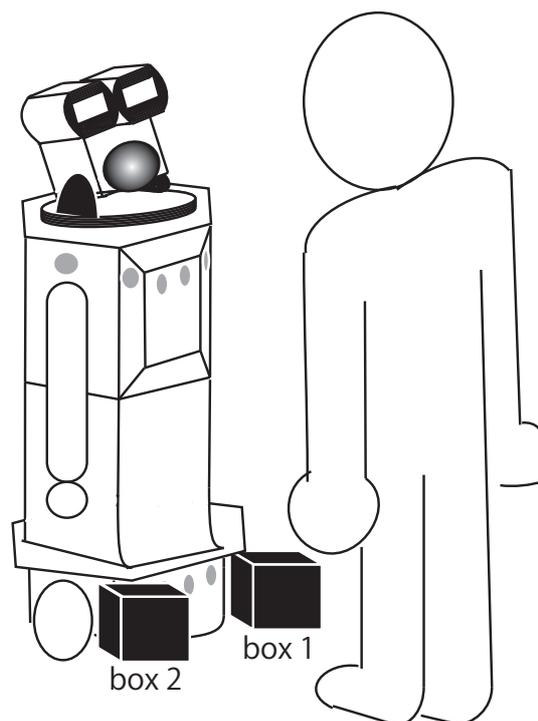


図1: 実験環境

ピーカは実験条件によって取り外すことはしない. よって実験条件に関わらず被験者は、Robovie の頭部のスピーカを目視可能である.

3.4 被験者

被験者は男女、合計52名でそのうち実験条件1に27名、実験条件2に25名にランダムに配分した.

3.5 実験の手順

実験では Robovie が箱2を被験者によけてもらうために、箱2へ視線(頭部)を向ける動作を繰り返しながら発話を行い、被験者が箱2をよけるかどうかの観察を行う. 実験の手順は、次の Step i から Step viii の流れに従う.

Step i Robovie の前に2つ箱を置き、首と手足を動かした状態で待機する(図2(a)).

Step ii 被験者が正面に立ったら、Robovie の顔の向きを被験者に向ける(図2(b)).

Step iii 頭部を動かし、Robovie の視線を箱2の方へ向ける(図2(c)).

Step iv Robovie が「KONOHAKO DOKETEKUDASAI」(この箱、退けてください)と発話する.

Step v 頭部を動かし、Robovie の視線を正面の方(被験者の方向)へ向ける(図2(d)).

Step vi 被験者が箱2をどかすまで Step iii から Step v を繰り返す(図2(e)).

Step vii 被験者が箱2をどかしたら Robovie が「ARIGATOUZOZAIMASITA」(ありがとうございました)と発話する(図2(f)).

Step viii 被験者に Robovie の発話の聞き取りやすさのアンケートを行う。

以上の流れを実行する前に被験者にはロボットの前に立つこと、そしてロボットの発話を聞くことを教示した。

Robovie は Step i の形で待機している。Robovie の Step i の動作は Robovie が生きている印象を被験者に思わせるためである。そして被験者は実験者の教示に従い Robovie の正面に立つ (Step ii)。Step ii で Robovie は胴体についている超音波センサによって被験者が正面に立ったかどうかを判定する。Step iii から Step v で繰り返される Robovie の頭部の一連の動き (Robovie の視線が箱 2 の方へ向き、その後正面を向く) は被験者に箱 2 に注意を向けることを促すためのしぐさである。頭部の動きで被験者の注意を引き込む行動は [3] を参考にデザインした。また、物体を参照する際に Robovie は指差しを利用しない。実験の目的は、物体の参照における音声の音源の位置の効果を検証することであるので、物体の指示において強い効果をもつ指差しはあえて排除している。実験では Step viii で Robovie の聞き取りやすさを被験者に 7 段階評価で判断してもらった。

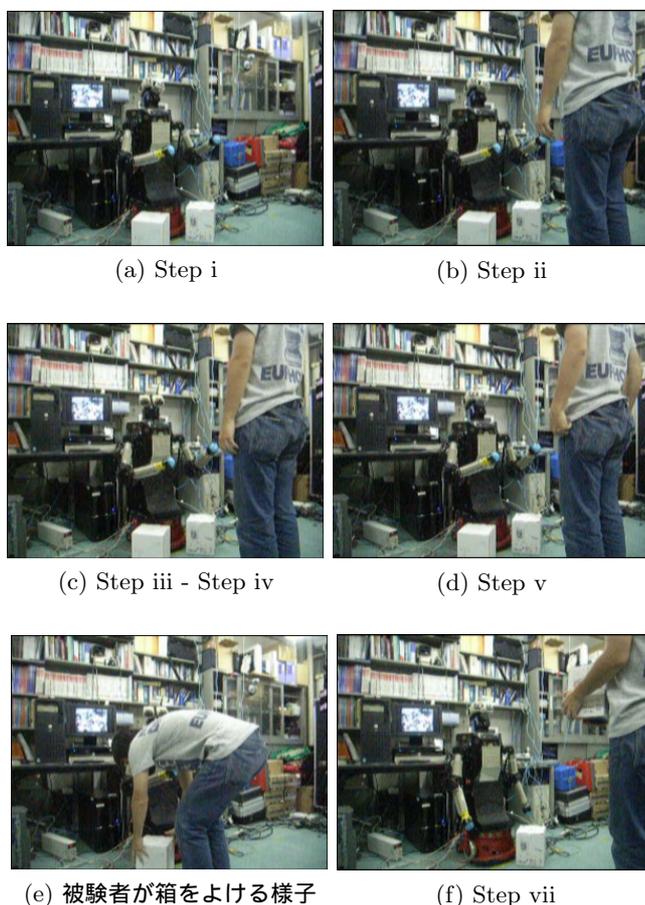


図 2: 実験手順

実験者は被験者が Robovie の発話を聞き取れているかどうか、またどの箱を動かすかを観察した。実験では 2 つの実験条件に対して Robovie の箱 2 をどけて欲しいことが被験者に正しく伝わるかを調べる。従って、正しく理解し箱 2 を動かした被験者と、箱 1 のみ、または両方の箱を動かした被験者に分ける形で被験者の行動を観察する。

3.6 実験結果の予測

仮説に照らし合わせ本実験で得られる結果を予測する。頭部のスピーカから発話音声が出力された場合、被験者は、頭部の向きに合わせて音声が箱 2 の方向に向けて出されていることを意識下で認識し、視線による物体参照と発話を理解して箱 2 のみを移動すると予想できる。一方で、腹部から発話音声が出力された音声が出力された場合、Robovie の頭部の方向とは無関係の方向 (被験者の方向) へ音声が出力されるため、被験者は、箱 2 のみを動かす被験者の数が減ることが予想できる。

3.7 実験結果

実験において、実験条件 1 と 2 の間でスピーカの聞き取りやすさに影響があったかの分析結果を示し、続いてどちらの箱をよけたのかについての結果を示す。

実験条件 1 と 2 のそれぞれについてロボットの発話音声の聞き取りやすさについて検証する。

実験条件 1, 2 それぞれについてロボットの発話を聞き取れたかを表 1 に示す。表 1 の結果を分析したところ、有意差は確認されなかった ($\chi^2(1) = 0.3611, p > .5$)。また発話の聞き取りやすさについて 7 段階評価による結果を分析したところ、両者の聞き取りやすさに有意な差は確認されなかった ($t = 0.2569, p > .7$)。以上の結果からスピーカを頭部に設置する場合と腹部に設置する場合においては発話の聞き取りやすさに相違はない。

実験条件 1 と実験条件 2 の被験者が箱 2 のみを移動させたかに関する比較結果を表 2 に示す。表 3 の結果を分析したところ、有意な差が確認された ($\chi^2(1) = 7.6453, p < .01$)。

本実験により頭部にスピーカを設置する方が腹部にスピーカを設置するよりロボットの物体の参照が人に伝わりやすいことが示された。なお、実験条件 1 と 2 をあわせて箱 2 のみを動かした被験者とその他の行動をとった被験者の間においてスピーカの聞き取りやすさに有意差はなかった ($t = 0.8232, p > 0.4175$)。以上の結果、本稿の仮説は検証された。

4. 考察

本章では、3 章で得られた実験結果を踏まえて考察を行う。

実験では Robovie のスピーカの位置を頭部に設置する場合と腹部に設置する場合の 2 条件を用意し、それぞれの条件において Robovie の物体を参照した発話が人に正しく伝わるかを検証した。実験条件の対照性を維持するために両実験とも頭部にスピーカが装着させている。人は映像を見ながら、その映像とは別の方向からその映像に関する音声を聞くとき、あたかもその音声が映像の方向から発せられているように錯覚することが知られている [5]。実験においても実験条件 2 ではロボットが腹部から音声を発しているが被験者からはロボットの音声が頭部に装着されたスピーカから発せられているものと錯覚し

表 1: スピーカの聞き取りやすさによる実験条件 1 と 2 の比較結果

	発話を聞き取れた人数	発話を聞き取れなかった人数
実験条件 1	26	1
実験条件 2	22	3

表 2: 箱の選択による実験条件 1 と 2 の比較結果

	箱 2 を移動	箱 1 を移動	両方の箱を移動
実験条件 1	14	0	12
実験条件 2	3	1	18

ていると考えられる。従って被験者から見たロボットの発話の様子は実験条件によって差が生じていないように思われる。

実験結果ではスピーカの位置を頭部に置く場合の方が腹部に置く場合よりもロボットの発話の参照対象が人に正しく伝わる事が確認された。両者の実験条件においてロボットの発話の聞き取りやすさに有意差、有意傾向はなかった。しかし、実験結果では両者の実験条件において被験者の物体の参照に有意差が確認された。従って、実世界の物体を参照する対話においてスピーカの位置の重要性が示された。

腹部のスピーカからロボットが発話する場合、腹部を動かさないで発話音声の方向は変化しない。一方で頭部のスピーカからロボットが発話する場合、ロボットが実世界の物体を参照する際にロボットの頭部は動かされるので発話音声の方向が変化する。従って人の物体参照に有意差が確認される理由に音源の方向の変化が挙げられる。頭部のスピーカからロボットが発話する方が実世界の物体を参照した対話に適していることは、音声の方向の変化が人によるロボットの視線による物体参照の認知に影響を与えていると考えられる。

5. まとめ

本稿では、ロボットのスピーカの位置が頭部にある場合と腹部にある場合で人への情報の伝達に影響を与えることを、実験により検証・確認した。

実験で用いたロボットは、頭部を動かし物体を参照する。頭部から音声を出力する場合には、頭部の動きに合わせて音声の出力方向が変化する。一方で腹部の場合は出力方向は変化しない。実験の結果、スピーカの位置は腹部よりも頭部に設置する方が被験者が正しく参照物体を理解することが確認された。

参考文献

- [1] V. Drygajlo, P. J. Prodanov, G. Ramel, M. Meisser, and R. Siegwart, "On developing a voice-enable interactive tour-guide robots," *Advanced Robotics*, vol.17, no.7, pp.599-616, Jan. 2003.
- [2] M. Fujita, "On activating human communications with pet-type robot AIBO," *Proc. IEEE*, vol.92, no.11, pp.1804-1813, Nov. 2004.
- [3] M. Imai, T. Ono and H. Ishiguro, "Physical relation and expression: Joint attention for human-robot Interaction," *IEEE Transaction on Industrial Electronics*, vol. 50, no. 4, pp. 636-643, Aug, 2003.
- [4] T. Kanda, H. Ishiguro, T. Ono, M. Imai and R. Nakatsu, "Development and Evaluation of an Interactive Humanoid Robot "Robovie ", " *IEEE International Conference on Robotics and Automation (ICRA 2002)*, pp.1848-1855, 2002.
- [5] G. H. Recanzone, "Auditory Influences on Visual Temporal Rate Perception," *J.Neurophysiol*, vol. 89, pp. 1078-1093, Feb 2003
- [6] K. Shinozawa, F. Naya, J. Yamato, and K.Kogure, "Differences in effect of robot and screen agent recommendations on human decision-making," *Int. J. of Human-Computer Studies*, pp.267-279, Dec. 2005.
- [7] K. Wada, and T. Shibata, "Living With Seal Robots - Its Sociopsychological and Physiological Influences on the Elderly at a Care House," *IEEE TRANSACTIONS ON ROBOTICS*, vol.23, no.5, pp.972-980, October 2007.
- [8] D. McNeill 著, 鹿取 廣人, 重野 純, 中越 佐智子, 溝渕 淳 共訳, 心理言語学「ことばと心」への新しいアプローチ, サイエンス社, 東京, 1990.
- [9] 石黒 浩, 宮下 敬宏, 神田 崇行, コミュニケーションロボット, オーム社, 東京, 2005.
- [10] 今井 倫太, 小野 哲雄, 石黒 浩, 中津 良平, 安西 祐一郎, "ロボットからの発話:自発的発話生成のための注意の表出機構の実現" *情報処理学会*, vol. 42, no. 11, pp.2618-2629, Nov. 2001.
- [11] 杉山 治, 神田 崇行, 今井 倫太, 石黒 浩, 萩田 紀博, 安西 祐一郎, "コミュニケーションロボットのための指さしと指示語を持ちいた 3 段階注意誘導モデル," *日本ロボット学会誌*, vol. 24, no. 8, pp.964-975, .