

# ロボットの動作によって人間の発話開始時間を制御できるか？

中山貴之<sup>†</sup> 廣井富<sup>††</sup> 黒田尚孝<sup>†</sup>  
三宅真司<sup>†††</sup> 伊藤彰則<sup>†††</sup>

本稿の目的は、人間とロボットの対話を円滑に行わせるために、ロボットの動作によって人間の発話開始時間を制御可能か検証することにある。実験として、ロボットが発話を終了する際に異なる開始時間でロボットに動作を行わせ、人間の発話開始時間を測定した。結果として、ロボットアバタの動作するタイミングを遅らせると実験参加者の発話開始時間も遅れることがわかった。

## Controlling the start time of human utterance by behavior of the robot

TAKAYUKI NAKAYAMA<sup>†</sup> YUTAKA HIROI<sup>††</sup> HISANORI KURODA<sup>†</sup>  
SHINJI MIYAKE<sup>†††</sup> AKINORI ITO<sup>†††</sup>

The purpose of this paper is to investigate possibility of controlling timing of human utterance by changing the behavior of the robot for smooth dialogue between human and robot. In this experiment, we examined different robot behaviors with different start time after the robot's utterance and measured the duration until the human utterance. As a result, it was suggested that the utterance start time of the participants was delayed when the timing delayed with the delayed robot behavior.

### 1. はじめに

筆者らは、日常生活を支援するロボットの研究・開発を行っている[1, 2]。例えば、重い荷物を人間に代わって運ぶことや指示された飲み物を運ぶことで人間の生活を支援しようという考えである。この様なロボットは、ロボットを扱うにあたり、特別な知識を有せず、操作できることが望ましい。そこで筆者らは、音声認識を用いてロボットに命令を与えている。近年、音声認識技術の進歩は目覚ましい[3]が、ロボットへの実装という観点からは課題が多い。音声認識を用いてロボットに命令を与えるためには、音声の認識精度の向上だけでなく、人間とロボットとのインタラクションをうまくデザインすることが重要である。筆者らは、ロボットに対応した音声認識精度向上のための音声認識手法の開発[4]、小型コミュニケーションロボット（ロボットアバタ）を用いたロボットデザイン[5]、ロボットデザインによる対話の容易さの調査[6]などを行ってきた。

人間同士の対話の場合は、言語情報だけでなく、非言語情報も参考にし、お互いの対話の間合いを取る。しかし、相手がロボットの場合には、対話の間合いを図るためのキューがうまく表出できないため、人間はいつ対話を開始すればよいか分からない。その結果、いつまでも対話が開始できない場合や、バグイン発話などの問題が生じる。

バグインについては、音響処理によりロボット発話をキャンセルする手法も用いられているが[7,8]、この方法を使えばバグイン発話をうまく認識することはできても、対話のタイミングのよさについての印象を向上させることは難しい。ロボットのジェスチャによって多人数会話のキューを表出する試みもあるが[9]、必ずしも個々の対話の印象を改善するものではなかった。

高杉ら[10]は、ロボットの発話開始タイミングおよびうなずき開始タイミングが対話にどのような影響を及ぼすかを調査した。これは、人間が先に発話した内容について4種類の応答時間を用意し、その影響を測る実験である。結果として、一部例外はあるが、応答時間が早いほど評価が高くなる傾向が見られたと報告している。また、うなずきに関しては、一部例外はあるが、ロボットのうなずき開始タイミングが発話開始タイミングと一致するとき評価が高くなる傾向が見られたと報告している。しかし、音声認識を行うことを考えると、応答時間を短くするには限界がある。そこで、筆者らは、ロボットの動作によって、人間の発話開始時間を制御可能かもしれないと考えた。本研究の目的は、ロボットが動作するタイミングを変えた場合に、人間の応答時間がどの様に影響を受けるか調査することである。これが明らかになると、音声認識精度の向上が見込まれることと、対話のタイミングのよさについての印象の向上が期待される。

### 2. ロボットアバタとは

人の日常生活を支援するロボットは、産業用ロボットの

<sup>†</sup> 大阪工業大学大学院 工学研究科 機械工学専攻  
Osaka Institute of Technology  
<sup>††</sup> 大阪工業大学 工学部 ロボット工学科  
Osaka Institute of Technology  
<sup>†††</sup> 東北大学大学院 工学研究科 通信工学専攻  
Tohoku University

ように単に性能が高いというだけでなく、人に対して違和感や恐怖感を与えないという親和性への配慮も必要である[11]. これらの解決法として、廣井ら[5]は、ロボットアバタという概念を提案した. 図1に示すように、作業を行う大ロボットの上にコミュニケーションを担当する小型のコミュニケーションロボットを配置し、この小型のコミュニケーションロボットの振る舞いによって、ロボット全体の親和性の向上を図るものである. この小型のコミュニケーションをロボットをロボットアバタと呼ぶ. 本研究では、このロボットアバタが搭載された日常生活支援ロボット Goyane [12]を用い、実験を行う.

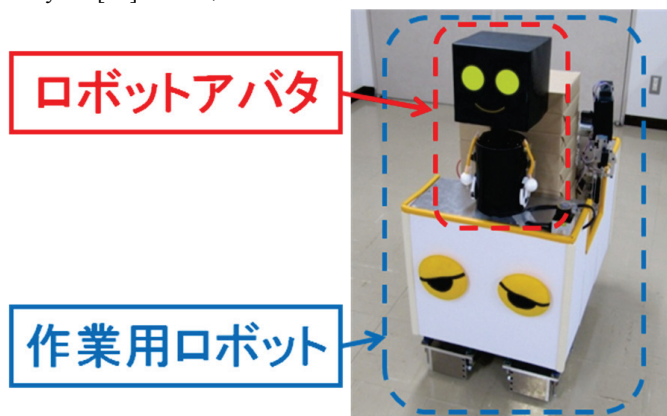


図1 ロボットアバタが搭載された日常生活支援ロボット Goyane

### 3. 実験

#### 3.1 実験目的

本実験の目的は、ロボットの動作によって、人間の発話開始時間を制御可能か調査することである. そこで、ロボットが人間に簡単な質問を行い、人間がそれに回答するという実験を計画した. ロボットは発話を開始し、ある一定時間が経過したのち、実験参加者の方へ顔を向ける、あるいは顔を下向きに下ろすという動作をさせる. この経過時間について、2種類の時間を設定した.

#### 3.2 実験環境

実験は、5000mm×8000mmの部屋で行った. 外乱光の影響をなくすため、窓のブラインドは閉め、ドアの小窓にはカーテンを施した. 実験参加者は20歳～23歳の男子大学生、大学院生18名である.

#### 3.3 実験装置

本実験は、前述した日常生活支援ロボット Goyane を使用した. 本実験では、ロボットアバタの目までの高さを1276mmに設定した[6]. また、実験参加者とロボットアバタとの顔までの距離を600mmと設定した.

ロボットアバタは、首や腕、胴体に自由度があるが、本実験で使用するロボットアバタは、図4(a)に示すように顔

に1自由度だけ有する実験専用として製作した. 外観を図4(b)に示す. 顔の角度は±50度動作可能とし、人に対し、視線方向を合わせることが可能にした. ロボットアバタの動作範囲について図5に示す. 首の制御は、アクチュエータとして Futaba 社製の RS405CB を用いた.

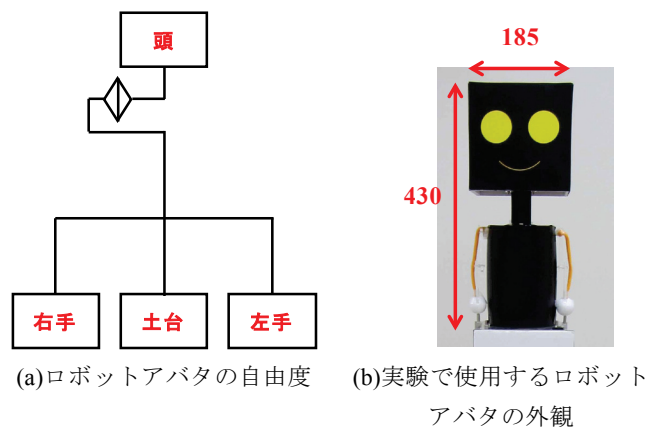


図4 ロボットアバタ



図5 ロボットアバタの頭部の動作範囲

対話システムは、実験参加者に対話自体に特別な印象を持たせないために、性別を尋ねて実験参加者に回答させるシステムとした. 音声の収録にはヘッドセットを用い、音声認識には Julius、音声合成には AquesTalk2 を使用した.

本実験では3台のノートPC (Panasonic 社製:Let's note CF-S9)を用いた. ノートPCの役割はロボットアバタ制御用、音声認識用、高さ変更機構制御用である. ロボットと実験参加者の位置関係を図6に、本実験の全体システムを図7に示す.

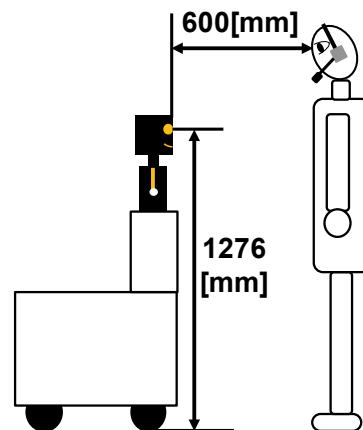


図6 ロボットと実験参加者の位置関係

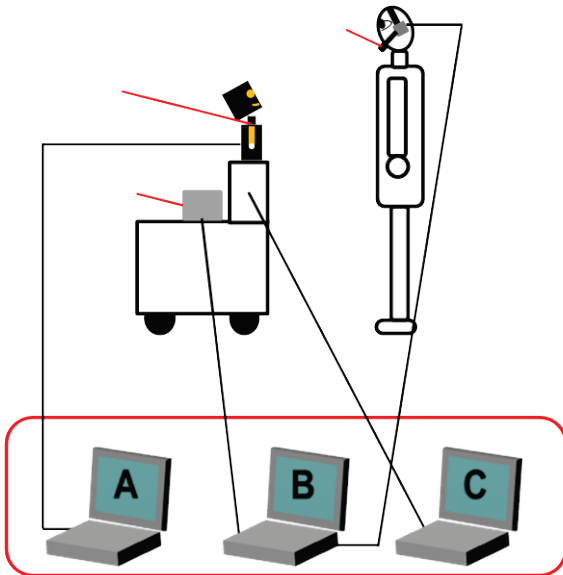


図7 システム構成図

### 3.4 実験内容

我々は、2つの仮説を立てた。一つ目は、「ロボットアバタの動作するタイミングを遅らせると実験参加者の発話開始時間も遅れる」、二つ目は「視線を向けた方が視線を外すよりも人は早く応答を開始する」である。一つ目を実験1と実験2とし、それぞれに対し、視線を人に向ける条件と視線を外す（下を向く）条件を設定した。

#### (1) 実験1

実験開始時にロボットアバタが自己紹介し、1回うなずく動作を行う。これは、うなずきを行うことで実験参加者の注意を引くと考えたためである[13]。図8に実験1のタイムチャートを示す。次にロボットアバタが実験参加者に性別を尋ねる。ロボットアバタが性別を尋ね終わる1s前からロボットアバタは動作を始め、実験条件の角度まで2sかけて頭部動作を行う。

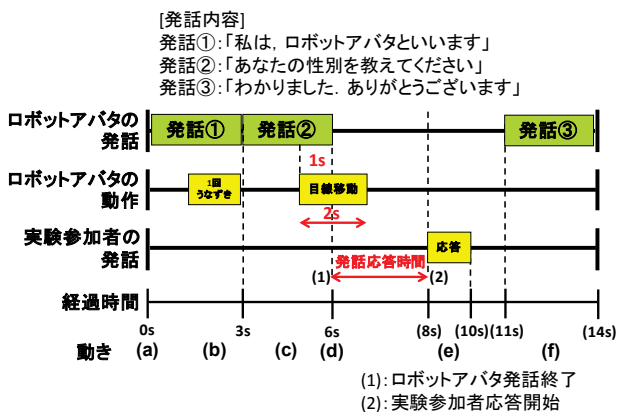


図8 実験1のタイムチャート

次に、二つ目の仮説である「視線を向けた方が視線を外すよりも人は早く応答を開始する」を検証するための実験条件について説明する。実験条件の内容として、ロボットアバタが質問を行った後、実験参加者に視線を向けることと、視線を外す（下を向く）の2種類の頭部動作をさせる。なお、ロボットアバタは実験開始時には正面を向いている。

#### ・実験条件 A

実験条件 A は、ロボットアバタが実験参加者に視線を向ける動作を行う。ここで本研究における「視線を向ける」について定義する。ロボットアバタには、眼球があるわけではないので、ここではロボットの顔の向きが実験参加者を向くことを意味する。実験参加者によって目の高さは異なるので、実験前に実験参加者の目の高さを測定した。このデータを用いて、図9に示すように角度  $\theta$  を下記式から算出した。  $H_e$  は床から実験参加者の目線までの高さ。  $A_e$  は、床面からロボットアバタの目線までの高さを示す。この角度を用いてロボットアバタの首の回転角度を制御した。

$$L = 600[mm]$$

$$X = H_e - A_e$$

$$\theta = \tan^{-1} \frac{X}{L}$$

実験条件 A の動作の様子を図10に示す。

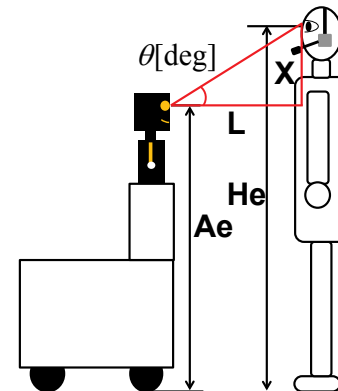


図9 視線制御

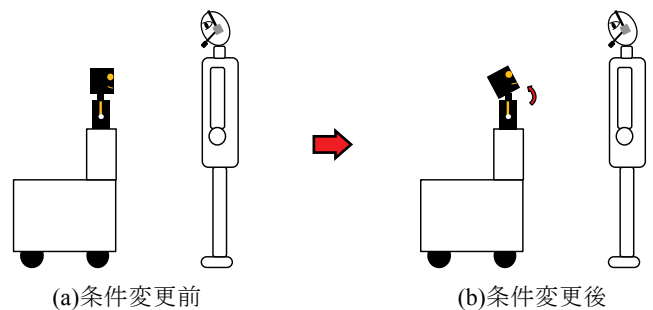


図10 実験条件 A

#### ・実験条件 B

実験条件 B はロボットアバタが視線を外す（下を向く）動作を行う。このときの角度は実験参加者とロボットアバ

タの目線が合わないよう<sup>1</sup>に 30deg（全ての実験参加者に対し、同一）に設定した。実験条件 B の動作の様子を図 11 に示す。

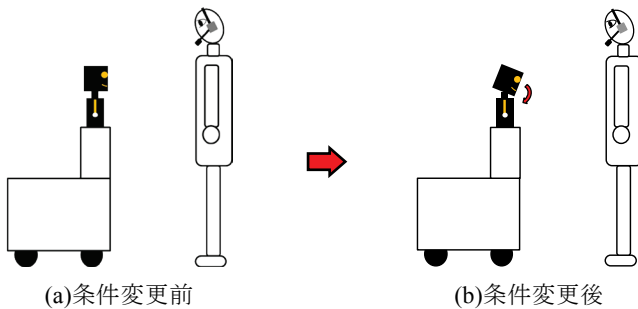


図 11 実験条件 B

以上のまとめとして、実験 1 におけるロボットアバタの動作を図 12 に示す。

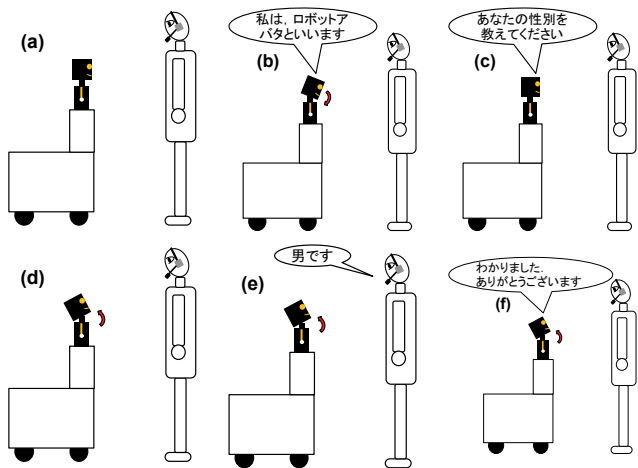


図 12 実験 1 のロボットアバタの動き

## (2) 実験 2

実験 2 は、実験 1 と比較しロボットアバタの頭部動作のタイミングを 0.5s 遅らせた。発話内容や実験条件については実験 1 と同様である。実験 2 の流れについてのタイムチャートを図 13 に、それに対応するロボットアバタの動作はタイミングが実験 1 と異なるだけであるため図を省略する。

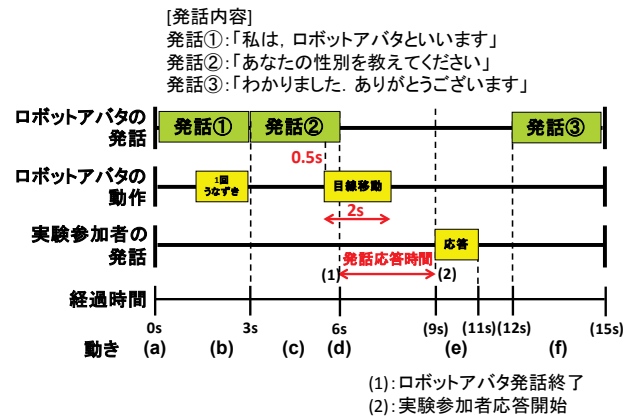


図 13 実験 2 のタイムチャート

## 3.5 実験手順

実験手順を以下に示す。

- (1) 実験参加者に実験を行う上での注意事項、実験の意図、実験内容を説明する。実験参加者が内容を理解し、同意が得られた場合は、同意書にその旨を記入してもらう
- (2) 実験参加者の身長と目の高さを測定する。
- (3) 治具を用いて実験参加者とロボットアバタの距離を 600mm に合わせる。
- (4) 実験参加者にヘッドセットを装着してもらい、音声対話について説明を行う。
- (5) 2 種類の実験条件の一方で実験を行う。
- (6) 音声対話を行う。

対話内容について以下に示す。

ロボットアバタ：「私は、ロボットアバタと申します。」  
 「あなたの性別を教えてください。」  
 実験参加者：「男/女です。」「男」「女」など  
 ロボットアバタ：「わかりました。」  
 「ありがとうございます。」

- (7) 実験終了後、実験参加者にアンケート用紙を記入してもらう。

アンケート内容に関して以下に示す。

「なぜ、あなたはそのタイミングでロボットに対し、応答を始めたか？」  
 アンケートに関しては実験が終わるたびに回答してもらった。

実験の順番については、実験参加者にランダムに設定した。条件については、視線を向けるまたは外すことのどちらか一方を実験参加者にランダムに設定した。なお、実験の様子はビデオで撮影した。

実験の様子を図 14 に示す。

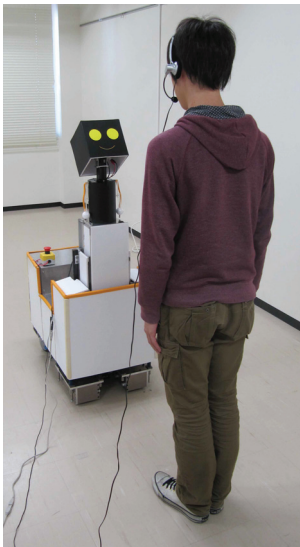


図 14 発話遅延実験（視線を向けている条件）

#### 4. 実験結果

ロボットの動作のタイミングを2種類用意し、人間の発話開始時間を測定した。さらに実験1と実験2の差（発話開始時間の差）について求めた。結果を表1に示す。次にそれぞれの実験の中で視線を向ける、外すという2種類の条件で実験を行った結果を表2に示す。ここでは、平均値のみを示す。

表1 実験1と実験2の比較

	実験1[s]	実験2[s]	実験1と2の差[s]
標準偏差	0.62	0.70	
平均	1.52	2.03	0.51

表2 実験条件別の実験結果

		実験1[s]	実験2[s]	実験1と2の差[s]
条件A	標準偏差	0.63	0.59	
	平均	1.47	2.10	0.63
条件B	標準偏差	0.61	0.78	
	平均	1.55	1.98	0.43

表1の実験1（ロボットアバタが発話終了の1s前から動作を開始する）と実験2（ロボットアバタが発話終了の0.5s前から動作を開始する）についての実験参加者の発話開始時間の差を検定する。つまり、「ロボットアバタの動作するタイミングを遅らせると実験参加者の発話開始時間も遅れ

る」を検証する。この2つの実験に対し、対応のあるt検定を行うと、有意水準1%以下で有意差が認められた。また、条件が異なる2群においてそれぞれ実験1と実験2について対応のあるt検定を行った結果、同様に有意差が認められた。よって、ロボットアバタの動作するタイミングを遅らせると、実験参加者の発話開始時間も遅れることが得られた。

次に「視線を向けた方が視線を外すよりも人は早く応答を開始する」について検証する。実験ごとに条件1と条件2について対応のないt検定を行った結果、有意差は認められなかった。つまり、視線を向けた方が視線を外すよりも人は早く応答を開始するとは言い難い結果が得られた。

#### 5. 考察

実験結果より、「ロボットアバタの動作するタイミングを遅らせると実験参加者の発話開始時間も遅れる」ことが得られた。実験1と実験2における、ロボットアバタの動作終了時刻の差は0.5sである。表1で示したように、実験1と実験2の差（発話開始時間の差）も同様に約0.5sであり、興味深い。撮影した動画を観察すると、実験参加者がロボットアバタの動作に注目し、動作が終了したのを待っていたと見ることができる。

2つ目の仮説である「視線を向けた方が視線を外すよりも人は早く応答を開始する」については、棄却された。横山ら[14]は、対話の流れの中で視線制御が対話に与える影響について、CGロボットと実ロボットにおいて比較した結果、実ロボットの方がCGロボットより視線制御の与える影響が大きいことを示した。これにより、実ロボットにおいて視線制御が有効であると思われたが、横山らは視線を向けない（本研究における視線を外す）条件について実験を行っていない。つまり、CGより実ロボットの方が視線制御が有効であることは示されていたが、実ロボットの視線制御が有効かどうかは示されていなかったと見こともできる。山本ら[15]は人間同士の指示・応答対話の分析を行い、武藤ら[16]は、人とアバタ（本研究とは異なるCGのアバタ）について同様の実験を行った結果、人とアバタの方が発話交代の時間を要した。このように、人同士と異なり、人とロボット間の対話では、人同士の常識がそのまま適用できない可能性が示唆された。

#### 6. 結論

本研究ではロボットが動作した場合に発話に対し、人間の応答時間がどの様に影響を及ぼすか調査した。「ロボットアバタの動作するタイミングを遅らせると実験参加者の発話開始時間も遅れる」可能性が示唆された。ロボットアバタが視線を向けても、外しても、発話開始時間に差は見ら

れなかった。これは、人同士の対話と人とロボットの対話  
においての違いと推測され、ロボットにはロボットに適し  
た動作が要求されることが示唆された。本研究の知見を活  
かし、ロボットとの対話のしやすさについて実験を行って  
いきたい。

美博, "音声対話インタフェースにおける発話タイミング制御と  
その評価," ヒューマンインタフェースシンポジウム 2007 講演  
会予稿集, pp.639-642 (2007)

## 参考文献

- 1) 廣井富, 黒田尚孝, 内藤圭祐, 高田晶太, 松井一馬, 井上駿,  
林和孝, 中山貴之, 松中翔平, 伊藤彰則, 日常生活支援移動ロボ  
ット ASAHI の開発-全体構想とハードウェア構成-, ロボティク  
ス・メカトロニクス講演会 2012 講演論文集 DVD-ROM,  
2A1-P03(2012)
- 2) 廣井富, 後藤基允, 山本祐三, 山根佑介, 稲田遥一, 大原達哉,  
木村昭太, 久野修平, 伊藤彰則, 日常生活支援移動ロボット  
Goyane の開発, ロボティクス・メカトロニクス講演会 2010 講  
演論文集 CDROM, 2A2-A21(2010)
- 3) J. Mariani Ed. "Spoken Language Processing", John Wiley &  
Sons (2009)
- 4) 伊藤彰則, 金山高志, 鈴木基之, 牧野正三: 小型ロボットに  
よる音声認識のための内部雑音抑圧法, ヒューマンインタフェ  
ース論文誌, Vol. 10, pp. 1-10 (2008)
- 5) 廣井富, 伊藤彰則, 中野栄二: 人間共存型ロボットのための  
ロボットアバタを用いた親しみ感の向上, 感性工学研究論文集,  
Vol.7, No.4, pp.797-805 (2008)
- 6) Y. Hiroi, T. Nakayama, H. Kuroda, S. Miyake and A. Ito, "Effect  
of Robot Height on Comfortableness of Spoken Dialog", Proc. of 5th  
Int. Conf. on Human System Interaction (2012)
- 7) 武田龍, 中臺一博, 高橋徹, 駒谷和範, 尾形哲也, 奥乃博:  
残響下でのバージン発話認識のための多入力独立成分分析を  
応用したロボット聴覚, 日本ロボット学会誌 Vol. 27, No. 7, pp.  
782-792 (2009)
- 8) 武田龍, 中臺一博, 高橋徹, 駒谷和範, 尾形哲也, 奥乃博:  
ロボット音声対話における Semi-blind ICA を用いた自己発話キ  
ャンセル, 情報処理学会全国大会講演論文集 6J-7, Vol. 5, pp. 27-28  
(2010)
- 9) B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro and N. Hagita,  
"Footing In Human-Robot Conversations: How Robots Might Shape  
Participant Roles Using Gaze Cues", Proc. of the 4th Int. Conf. on  
Human Robot Interaction, pp. 61-68 (2009)
- 10) 高杉將司, 吉田祥平, 沖津健吾, 横山正典, 山本知仁, 三宅  
美博: コミュニケーションロボットとの対話における交替潜時長  
と領き先行時間長の影響評価, 計測自動制御学会論文集, Vol.46,  
No.1, pp.72-81 (2010)
- 11) Y. Hiroi and A. Ito, "Effect of the Size Factor on Psychological  
Threat of a Mobile Robot Moving Toward Human", KANSEI  
Engineering International, vol. 8, no. 1, pp.51-58(2009)
- 12) 廣井富, 篠原達也, 兼次一喜, 岩本昂, 中山貴之, 伊藤彰則:  
日常生活支援移動ロボット Goyane の開発 -高さ変更可能な機構  
の提案-, ロボティクス・メカトロニクス講演会 2011 講演論文集  
DVD, 2P2-Q08 (2011)
- 13) 渡辺富夫, 大久保雅史, 小川浩基: 発話音声に基づく身体  
的インタラクションロボットシステム, 日本機械学会論文集(C  
編)66 巻 648 号, pp.251-258 (2000)
- 14) 横山真男, 青山一美, 菊池英明, 帆足啓一郎, 白井克彦: 人間  
型ロボットの対話インターフェースにおける発話交替時の非言  
語情報の制御, 情報処理学会論文誌, Vol.40, No.2. pp.487-497  
(1999)
- 15) 山本知仁, 平野作美, 小林洋平, 高野弘二, 武藤ゆみ子, 三宅  
美博, "対話コミュニケーションにおける 2 種類の発話タイミン  
グ相関," ヒューマンインタフェースシンポジウム 2007 講演会  
予稿集, pp.631-634 (2007)
- 16) 武藤ゆみ子, 高野弘二, 大良宏樹, 小林洋平, 山本知仁, 三宅