

Talking-Ally における宛名性の表示機構について

Intended to Integrate Mechanism of ‘Addressivity’ for Talking-Ally

蔵田 洋平^{1*} 小田原 雄紀¹ 松下 仁美¹ 大島 直樹¹
P. Ravindra S. De Silva¹ 岡田 美智男¹

Yohei Kurata¹, Yuki Odahara¹, Hitomi Matsushita¹, Naoki Ohshima¹,
P. Ravindra S. De Silva¹ and Michio Okada¹

¹ 豊橋技術科学大学 情報・知能工学系

¹Department of Computer Science and Engineering,
Toyohashi University of Technology

Abstract: In our daily conversation, speaker is always referring addresser’s behaviors to construct and adjust his/her conversation. We intended to propose the similar kind of potentially for Talking-Ally by integrating the concept of “Addressivity”. Talking-Ally is employed a novel utterance generation mechanism by considering the user’s eye-gaze behaviors and attention, these mechanism is convey that robot is constructing its conversation toward interactive person. In this paper, we introduced the design concept of robot and utterance generation mechanism.

1 はじめに

私たちが発話を行うとき、その発話に向ける相手がいるだろう。この相手が、私たちの発話を受け取り、応答する聞き手となってくれるため、私たちは発話を行うことができる。このとき、相手の興味が別の場所に向けられている場合や、発話に気づいてないと、その発話は宙に浮いてしまい、意味のないものになってしまう。つまり、私たちが話し手であるのは、相手が聞き手になってくれているという事実を支えられている[1]。そのような聞き手に対し、話し手は、聞き手の状態に合わせて発話のアクセントやスピード、内容を調整し、また相手の方を向き、視線を合わせるといった身体動作を交えながら発話を行う。このように、聞き手の状態に合わせて、聞き手の方に向けた発話には、「宛名」が備わっているといえるだろう。

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、聞き手の状態に関わらず、同じような発話を機械的に繰り返していた。そのような発話は、私たちにとって自分に向けられたものだと感じられず、応答責任を引き出されないため、簡単に無視できるものとみなしてしまう。その意味で、これまでの発話生成システムや音声合成システムは宛名を欠いていた。また、聞き手がその相手が関わりを持とうとしないうという意味で、聞き手の発話を組織する際の聞き手



図 1: Talking-Ally

の参加余地も奪っていた。このように、これまでの音声合成システムは、聞き手を配慮せず、「宛名のない合成音」の聴取を一方的に強いられてきた。

そこで本研究では、聞き手の状態を意識しながら発話を組織し、聞き手に対し志向を表示する発話生成システム、Talking-Ally を構築した。以下、研究背景およびシステム概要等について述べる。

2 研究の背景

話し手が発話を行うとき、その発話を言葉や顔の向き、視線、身振りといった身体的動作によって、聞き手に向ける。聞き手は、そのような話し手の様子から、話し手に対し応答を返す。話し手はその応答をリソー

*連絡先：豊橋技術科学大学 情報・知能工学系
〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘1-1
E-mail: odahara@icd.cs.tut.ac.jp

スに発話を調整する、といったように、会話は、話し手と聞き手が相互に調整し合うことによって組織される。以下では、話し手が聞き手に向ける志向について、話し手と聞き手が相互に調整し合うことによって組織された発話に伴ってくる宛名について議論する。

2.1 「宛名性 (addressivity)」

「言葉というのは、本来だれかに向けられたもので、必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、Bakhtin の「宛名性」に関する指摘である [2]。また Bakhtin は、話し手は必ず「いまこの発話を聞いてほしい相手」へと宛名を向け、その相手との関係や状況に合わせて適切に発話を構築するという。この意味で、発話には必ず「宛名」が伴っているべきだと考えられる。「宛名」を伴った発話は、聞き手の、会話への参加を促すことができる。ここで、話し手が相手へと志向を向ける際には、言葉の他に、身体や視線の向きにおいても重要な役割をもっている。このような志向について、以下で説明する。

2.1.1 身体方向による志向の表示

会話において話し手が、聞き手に志向して話しかける際に、言葉に伴って身体が相手の方向を向いているだろう。発話においては、このような身体の動きが言葉に随伴して起こり、発話を相手に伝える上で大きな役割を持つという [3]。

2.1.2 視線による志向の表示

Schefflen や Kendon の指摘によると、身体は志向の表示に関して階層化されており、最も基本的な志向を表示する腰の前面方向から、上体の前面方向、手、指あるいは顔、目の順でより先鋭な志向を表示する [4][5]。また、浜田は、「目には向き、つまり志向性があり、そこにはっきり相手が主体として私の方に向かって力をを感じる」という [6]。このように会話において話し手の視線は、志向を示す役割が特に大きいものとなっている。

2.2 聞き手との関係の中で生じる宛名

話し手は聞き手に宛名を向けようとする時、聞き手の示す「聞き手性 (hearerhood)」を常に意識しながら、発話や振る舞いを調整している。Goodwin によれば、「聞き手性」とは会話において、聞き手が話し手に向ける、「今、あなたの話を聞いていますよ」といった社会的表示のことである [7]。

話し手が、この「聞き手性」を考慮しながら発話を調整し、更に聞き手はその様子から、話し手に「聞き手性」を表示する。このような、話し手と聞き手との互いを意識しあう関係が形成された時、会話が組織さ

れたといえるだろう。会話の場では、話し手または聞き手の志向が少しの間逸れても、会話が途切れることはない。これは、関係によって組織された発話には、宛名が備わっているためだろう。

2.3 発話の「非流暢性」

普段の会話の中での発話は、ニュースキャスターがニュースを読み上げるような流暢な発話であることはほとんどない。自然な発話は、「えーと」「あのー」といった言い淀みや言い直しを含む非流暢なものが多いのではないだろうか。

このような非流暢な発話は、要因が話し手だけの働きにより生まれてものだと考えてしまっていた。一方、非流暢な発話は、聞き手との関わり合いの中で生まれてくるのではないかと考えられる。聞き手が話し手の発話に興味がないようなら、つい沈黙してしまったり、聞き手が理解できていないようなら、聞き手が理解しやすいよう、言い換えたりするだろう。このように、発話の非流暢さは、話し手が聞き手と発話を調整し合う結果として、現れてくる。

言い直しや言い淀みを含む非流暢な発話は、弱々しく、聞き手からアシストを引き出すことのできる、聞き手の会話への参加余地を残した発話になるのではないだろうか。このような発話は、聞き手にとって自分を意識し、調整された発話に聞こえ、発話に宛名を持っていると感じることができよう。

2.4 発話に対する志向的な構え

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、ただ作られた通りに動いていると感じる、設計的な構え (design stance)、その動きに意図や信念があると感じてしまう志向的な構え (intentional stance) 等のいずれかの構えを取ると指摘している [8]。これは発話生成システムにも当てはまるだろう。

例えば、平坦なスピードで、機械的に行われる発話に対して、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、志向的な構えを取るように思われる。私たちが、発話生成システムに志向的な構えを取ったとき、発話生成システムを意識した振る舞いをとるのではないだろうか。ユーザと発話生成システムが互いに意識し合うことによって、人らしくより自然な会話となると考えられる。

3 発話生成システム Talking-Ally

これまでに、聞き手の視線を捉えながら発話を組織する発話生成システムが聞き手の会話への参加態度を引き出すことに効果があることを確認した [9]。ここでは、

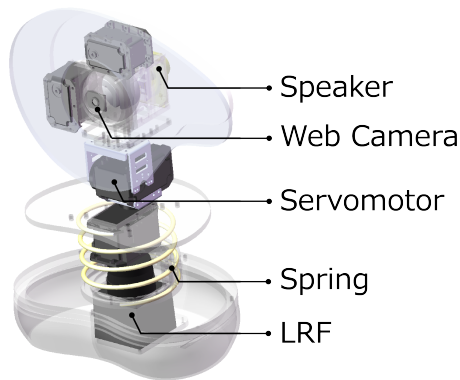


図 2: Talking-Ally のハードウェア構成

宛名についての議論を進めるために、発話生成システムに視線表出機構の追加を試み、新たな Talking-Ally を再構築した (図 2)。

3.1 志向表示機構

会話の中で話し手は、聞き手に対して志向しながら発話を行っている。聞き手に対して志向を向けるとき、話し手は言葉の他に、身体向きを用いる。Talking-Ally には、人に対し身体向きで志向を示すために、頭部機構、眼球機構を備えている。

3.1.1 頭部機構

会話において、発話を聞き手に向けるためには、身体向きが重要な要素となる。そのため、Talking-Ally は、頭部の動作させることによって、聞き手への志向を示す。2つのサーボモータによる2軸動作により、頭部の縦振り・横振りを可能としている (図 3)。この動きにより、聞き手の方向に顔を向けることができる。

3.1.2 眼球機構

視線は、聞き手に志向を示す役割が特に大きいと述べた。そこで、発話生成システムにおける視線の役割を確認するため、Talking-Ally には、視線表出機構を実装した。特徴的な「大きな一つ目」をもたせることで、視線による志向性を強調している。眼球機構は2つのサーボモータにより、上下左右に視線を移動させることが可能となっている (図 4)。

3.2 聞き手のセンシング

聞き手に対し、宛名を示すためには、「聞き手性」を考慮しながら発話を調整する必要があると述べた。このため Talking-Ally では、測域センサ (LRF) と、眼球内の Web カメラにより、聞き手の様子の把握を行い、発話内容・振る舞いの調整を行う。LRF は、センサ中心からのある物体の距離、角度がわかるため、人物の



図 3: 頭部機構



図 4: 眼球機構

位置検出に用いている。聞き手の距離も会話においては重要である。Hall は、人と人との間に、その関わりに応じて、親密距離、個体距離、社会距離、公衆距離などの対人距離が存在することを指摘した [10]。さらに西出は、対人距離において、50 cm ~ 1.5 m が普段の会話が行われる会話域としている [11]。本システムの会話のタイミングは、この会話域を基にした。また、眼球内の WEB カメラを用い、OpenCV による顔検出による、聞き手の検出も行っている。

3.3 Talking-Ally の発話

音声合成システムとして、ATR-P 社の「Wizard Voice」を採用した。これは、子供の声を発話する音声合成エンジンであり、その声は非常にかわいらしいものとなっている。これにより、聞き手に「弱い存在」として Talking-Ally を解釈させることで、聞き手のアシストを引き出すことを狙いとする。また、発話内容については聞き手との距離などの状態に応じて、発話内容に「えーと」「あの一」「それでね」といったターン開始要素 (Turn-initials) や、「ね」「だって」などの発話末モダリティの付与を行い、そして言い直しや言い淀みを動的に選択するといった、非流暢な発話を行う (図 5)。

3.4 デザイン

Talking-Ally のデザインは、手や足などの機構を設けないことで、足りない部分を人が思わずがアシストしてしまうような、積極的な振る舞いを引き出すことを狙いとしている。

また、バネを用いて、不規則で機構にとらわれないヨタヨタとした動きや、全体的に丸みを持った形状な

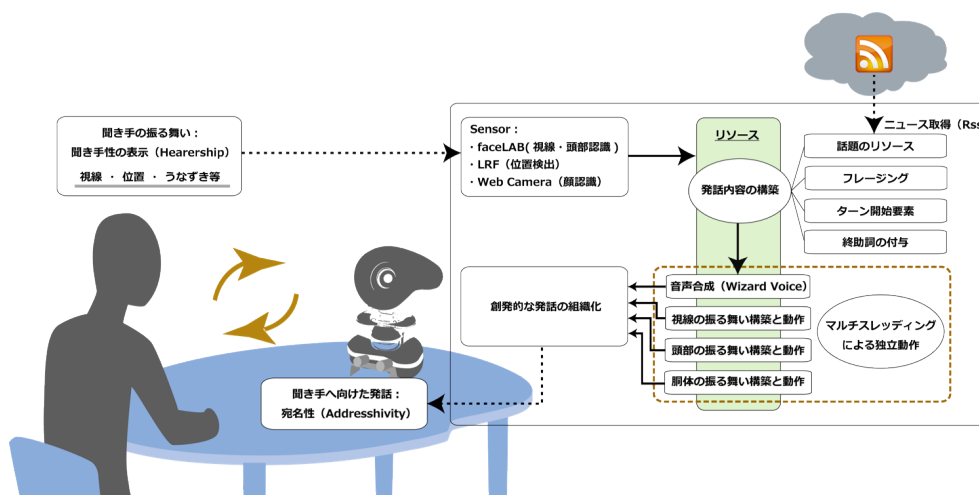


図 5: システム構成

どから、親しみを持ちやすい生き物のようなデザインとした。

4 動作例

Talking-Ally は、測域センサにより、周囲の状況を常に伺っている。この場に人が近づくと、Talking-Ally は人が近づいた場所に頭部と視線を向ける。この様子から、私たちは Talking-Ally が私たちに、興味があるかのように感じられる。また、Talking-Ally の様子を見続けていると、こちらに何か訴えかけたいようにも感じる。そのような Talking-Ally の様子が気になり、人が Talking-Ally に前述した対人距離における会話距離 (3.2 節) まで近づくと、Talking-Ally は発話を始める。Talking-Ally の発話内容は、RSS から取得した最新ニュースについての話題である。Talking-Ally は発話内容のフレーズごとに、「～ね」といったモダリティや、首を縦に振る動作を発話の際に付与する。また、聞き手の状態から、言いよんどんだり、沈黙したりする。こういった Talking-Ally の様子からは、こちらに対し、宛名を伴った発話を行っているように感じられた。

5 まとめと今後の展望

本研究では、発話に宛名性を持たせることを目的とする発話生成システム実現のため、視線表出機構に着目して、Talking-Ally を再構築した。測域センサと WEB カメラにより、人の距離などによって、発話や振る舞いを行わせている。また、聞き手に対する宛名性の表示として、Talking-Ally に自由に動く眼球を備え、視線による志向性の表示を可能とした。今後は、評価実験を行い、Talking-Ally の発話に与える影響について評価していきたい。

謝辞

本研究の一部は、科研費補助金 (挑戦的萌芽研究 24650053) の助成による。ここに記して、感謝の意を表する。

参考文献

- [1] 岡田美智男: 『弱いロボット』, シリーズ ケアをひらく, 医学書院 (2012).
- [2] ミハエル・バフチン (伊東 訳): 『小説の言葉』, 平凡社ライブラリー, 平凡社 (1996).
- [3] 山田誠二: 『人とロボットの＜間＞をデザインする』, 東京電機大学出版社 (2007).
- [4] Scheflen, A. E.: *Communicational Structure*, Bloomington: Indiana University Press (1973).
- [5] Kendon, A.: *Conducting Interaction*, Cambridge: Cambridge University Press (1990).
- [6] 浜田寿美男: 『「私」というものの成り立ち』, ミネルバ書房 (1992).
- [7] Goodwin, C.: *Conversational Organization: Interaction between speakers and hearers*, Academic Press (1981).
- [8] ダニエル・デネット (土屋俊訳): 『心はどこにあるのか』, サイエンスマスターズ, 草思社 (1997).
- [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ポジウム 2012 論文集, 2E-2, (2012)
- [10] エドワード・ホール, (日高 敏隆, 佐藤信行 訳): かくれた次元, みすず書房, (1970)
- [11] 西出和彦: 人と人との間の距離, 人間の心理・生態からの建築計画 (1), 建築と実務, No. 5, pp. 95–99, (1985).