

Talking-Ally : 聞き手性をリソースとする発話生成系の実現にむけて

Talking-Ally : Utterance Generation Mechanism by Collaborating with a Hearer

小田原 雄紀¹ 蔵田 洋平¹ 大島 直樹¹ P.Ravindra S. De Silva¹ 岡田 美智男¹
Yuki Odahara¹, Yohei Kurata¹, 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: The next generation social robotic might be utilized spontaneous speech is systematically shaped by planning the conversation, gaining attention from the addresser, error-handling and refreshing the conversation. Since, our Talking-Ally utilizes similar mechanism that considering the resource of addressivity and hearership by conjunction with turn-initial, and entrust behaviors to enhance the engagement of the conversation. In this study, we discuss about the concept of novel utterance generation mechanism and its performance in human-robot interactions.

1.はじめに

私たちは大地の上を歩くと同時に、大地が私たちが歩かせている。一度、そうした視点から私たちの身の回りのことを見渡してみるのも面白い。街の中をなにげなく歩く時に、「街を歩いているのは、この私だ」と自分を中心に考えてしまいがちだけれど、その歩道や街の看板、駅に向かう人の流れは、私たちの歩くという行為を制約し方向づける。その意味で、なにげない一歩とそれを支える大地や街並みとが一緒になって「歩行」という行為を生みだしているといえる[19].

これは「歩く」という行為に限られない。人前で何かを話そうとするとき、その相手が無表情なときには、とても話し難いということがある。なにげなく言葉を繰り出そうとするとき、その相手は私たちの発話が向かう対象であると同時に、私たちの発話の内容を制約し方向づける。つまり、その聞き手のうなずきや視線の動きに支えらながら、一緒になって発話を生みだしているといえる[20].

一方で、音声合成システムの場合はどうだろう。スイッチをオンにすると、そのスピーカーからは確かに合成音は聞こえてくる。しかし、その合成音は「宛名」を伴ったものではない。その発話は聞き手の存在を予定したものでも、その状態に配慮したも



図1 Talking-Allyの外観

のでもない。私たちがテキストの読み上げソフトを使用する際には、そうした「宛名のない合成音」の聴取を一方的に強いられてきたともいえる。

本研究では、私たちの発話はむしろ聞き手との相互行為的調整の結果として組織されたものと捉え直し、これらの観点から、聞き手の存在を予定しつつ、その聞き手との相互行為的調整に基づいて発話の組織化を行う新たな発話生成システムの構築を行った。以降の章では、本研究の背景となるポイントを整理するとともに、プロトタイプシステムの概要、評価結果の一部について述べる。

2. 研究の背景

2.1. 「オープンなシステム」としての身体

私たちは行為するとき、その行為を自分一人で生み出したものであると捉えてしまう傾向にある。これは、「私たちの身体は外から容易に観察できることから、個として完結しているような先入観を持たれやすい」ためだと考えられる[1]。これまでのロボットや発話生成システムの開発は、この「個体能力主義」的な考え方によるものが多い。観察できる行為全ての要因はその行為主体のみに帰属するとして、ロボットの機能を作り込み、その行為を実現しようと努力してきた。こうして作り込まれたロボットは想定された環境内では力を発揮できるが、実環境の動的で多様な変化の前では、その存在感を失ってしまう。

この流動的な環境に適応して行為を繰り返すためには、行為主体と環境の関係に視点をシフトする必要がある。私たちは自分一人では行為を完結することができないという「身体の不定性」を内包することで、環境に柔軟に対応し、環境との切り結びの中で結果として行為の意味を完結させる。その意味で私たちの身体は環境に対して開かれており、「オープンなシステム」といえる[2]。

このオープンなシステムとしての身体観は会話におけるコミュニケーションの分野に展開することが可能だと考えられる。話し手が「おはよう」と語りかけたとき、聞き手がその発話を無視すれば「おはよう」という言葉は、ただの意味の備わっていない音声として宙に浮いてしまう。また、話し手は聞き手の視線遷移などの多様に変化する振る舞いに柔軟に対応しながら、発話を繰り返している。これらから話し手の発話も不定性を内包しており、聞き手(環境)に支えられることによって、発話の意味を完結させていると考えられる。

2.2. 発話の「非流暢性 (disfluencies)」

私たちは行為するときその過程で起こる環境の変化に合わせて、微細な修正を施す。この微細な修正のときに起こるわずかな淀みのことを「マイクロスリップ(micro slips)」と呼ぶ[3]。「私たちの行為はほとんど気づかれることのない「わずかなスリップ(=マイクロスリップ)」とともに行われる傾向が強くなり、行為がマイクロにスリップできることで環境に対して「柔軟」であることが保証されているのである」という指摘から、「マイクロスリップ」は私たちの行為と流動的な環境との切り結びの中に生まれる「柔軟な調整過程」が顕在化したものだといえる。実際にインスタントコーヒーを入れる作業のような

日常的な行為において、こうした現象は多く観察される[4]。またテーブル上でコーヒーを入れる際に、不必要なものがテーブル上に多く配置されている状態など、環境が複雑化するとマイクロスリップの生起頻度が増加するという現象も観察されている。これらのマイクロスリップに関する議論は生態心理学の分野で盛んに行われている[5][6]。

この視点を発話生成、社会的相互行為に関して拡張すると、発話における「非流暢性 (disfluencies)」について新たな議論の展開が可能になると考えられる。発話の非流暢性とは、自然な発話(spontaneous speech)の中で繰り返される「あー」、「えっと」、「それでね」のような「言い淀み」や「言い直し」などを指す。一見、意味のないように感じるこれらのフレーズは、話し手が自分の話を開始するための指標として利用したり(ターン開始要素)[7][8]、聞き手に配慮し発話の調子を整えるための表示、話し手が発話を継続すること(発話権の保持)を示すため[9]など重要な役割を果たす。

これまで、これらの発話の非流暢性に関する言語心理学的な説明の多くは、その要因を発話のプランと行為との乖離(=スリップ)とその修復プロセスとして、「話し手」である発話者の内部のメカニズムにのみ帰属させる傾向が強かった[10]。

本研究では、発話の非流暢性を「話し手 (speaker)」と「聞き手 (hearer)」が相互に調整した結果として、生まれるものと仮定する。つまり、非流暢性を発話におけるマイクロスリップとして捉え直す。

2.3. 「宛名性 (addressivity)」

バフチンは「言葉というのは、本来は誰かに向けられたもので、必ず宛名を伴ったものだ」という[11]。この指摘から、話し手の発話は「いまこの瞬間に発話を受け取って欲しい聞き手」へ向けられており、必ず宛名を伴っていると考えられる。

では、自動販売機の「ありがとう」という発話に「宛名」は伴っているのだろうか。飲み物を買った人がどんな人であっても、自動販売機はいつでも同じような「ありがとう」を淡々と発話すると考えられる。聞き手へ向けた発話を行うためには、聞き手の様子を気にしながらその状況にふさわしい発話を組織する必要がある。その意味で、自動販売機の「ありがとう」には「宛名」が伴っていなかったと考えられる。同様に、これまでの発話生成システムを備えたロボットたちの発話も「宛名」を伴ったものではなかった。一方的に発話を繰り返す、聞き手をただ戸惑わせ、聞き手の発話への参加余地まで奪い去ってしまう。

発話に「宛名性 (addressivity)」を伴わせることができれば、聞き手に発話の組織化への参加余地を提

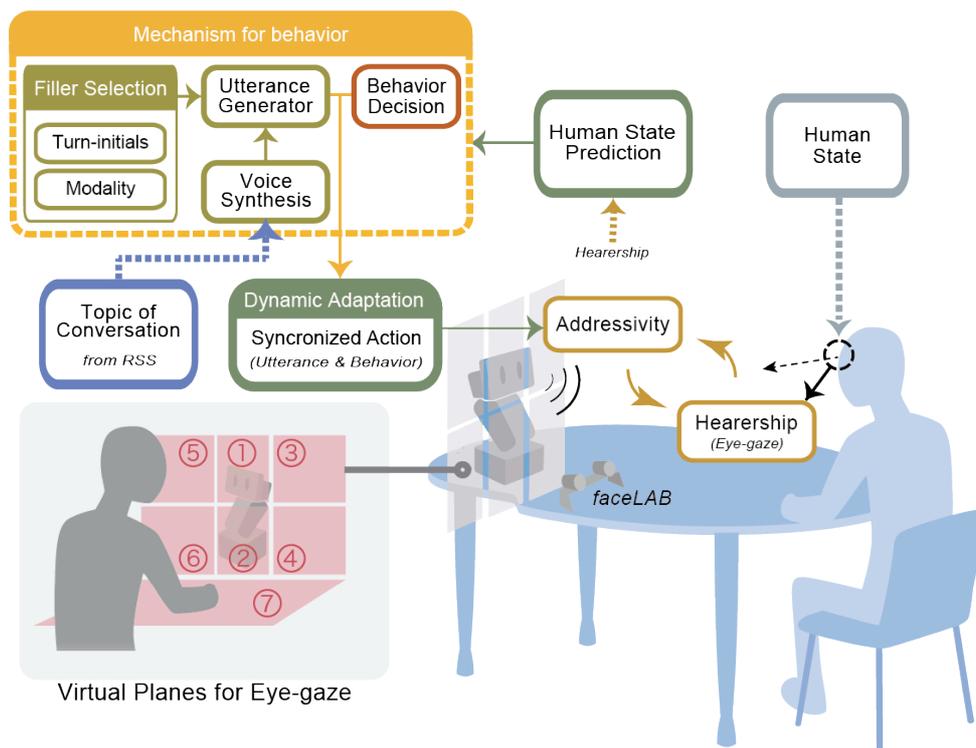


図2 Talking-Ally のシステム構成

供することが可能になると考えられる。これが聞き手と相互行為的に調整を行うことを可能とする一つの手掛かりとなり、聞き手を揺り動かすような力を備えた発話の組織化につながると考えている。

2.4. 「聞き手性 (hearer-ship)」

そのような発話を向けられたとき、私たちは思わず応答責任を感じる。そして発話が向けられた聞き手は、この応答責任に応じるべく視線、身振り手振り、うなずきなど様々な表示を用いて「いま、あなたの話をきいていますよ」ということを発話者である話し手へ伝える。

この視点を裏付けるものの一つとして Goodwin の「聞き手性 (hearer-ship)」に関する指摘がある[12]。Goodwin によれば「話し手は発話の冒頭で聞き手の視線を獲得できていないならば発話を一時中断し、聞き手の視線が得られた時点で発話を再開する」という振る舞いを観察できるとしている。

また、聞き手が話し手を見るという行為は会話に参加しようとする態度 (engagement) を示す一つの方法であるとし、もし参加者が互いに視線を合わさないう場合は、一時的に会話に参加しようとする態度 (disengagement) の表示になるという。このことから、視線は会話への参加態度を示す重要な

手掛かりの一つだといえる。

これをうけて、中井は会話への参加態度の示し方について、言語的・非言語的行動の関係を会話分析により明らかにしている[13]。この指摘によれば、言語行動に随伴させながら視線や身振り手振り、うなずきなどが起こり、会話の空間内でそれらが行われることによって参加態度を示すことにつながるとし、逆に会話の空間から外すことで、不参加態度を示すとしている。

これらから、聞き手の「話を聞いている」という「聞き手性」の表示を捉える (=「リソース」とする) ことは、聞き手の会話への参加態度や関心の度合いを知る手掛かりとなり、聞き手の状態を反映した発話を組織することにつながると考えられる。

2.5. 発話に対する志向的な構えと説得性

これまでの発話生成システムは、設計通りに発話する機械的なものとして、「設計的な構え」で捉えられる傾向が強かった。この設計的な構えによって認識されるような作り込まれた発話生成システムやロボットはいずれ、人々によって飽きられてしまう傾向にある[14]。社会的に受け入れられ、飽きることのない発話生成システムを構築するためには、聞き手の「志向的な構え (intentional stance)」[15]を引き

出す必要がある。「志向的な構え」とは、私たちが「主体的に行動しているモノ」をみたとき、「自発的に、合目的に意思決定をしている」と解釈してしまう帰属傾向のことをいう。本発話生成システムでは聞き手に配慮することで、このスタンスを聞き手から引き出すことを狙いとしている。そして、志向的な構えを引き出すことが可能な発話生成システムをロボットが備えたならば、会話というコミュニケーションを行う場において、主体的に発話を行う、合理的な活動体としてのロボットを生み出す一つの手掛かりになると考えている。

また、Fogg らによって盛んに議論されている「カプトロジ (CAPTOLOGY: Computer As Persuasive TechnOLOGY)」という分野[16]をベースとして、発話に聞き手を揺り動かすような「説得的な力」備えさせることも可能になると考えている。ミーティングなどの議論の場だけに限らず、私たちの日常的な会話は多くの場面で説得的だと考えられる。その意味で発話に「説得性」を備えるさせることは非常に有用だといえる。

これらが可能ならば、これまでの聞き手に一方的に支えてもらっていた発話生成システムから、聞き手とともに支える、聞き手に優しい発話生成システムへのシフトとその提供につながると考えている。

3. プロトタイプ的设计

3.1. Talking-Ally の構築

2 章での議論を考慮に入れた発話生成システムを実装するためのプラットフォームとして、図 1 に示す Talking-Ally を構築した。Talking-Ally は以下の特徴を持つ。

(a) かわいらしいデザイン

乳幼児や小動物などの「かわいらしい」存在に対して、私たちは親しみを感じたり、自分から積極的に関わりを持つとうとする。Talking-Ally をかわいらしいデザインのロボットとすることで、聞き手の積極的な関わりを引き出すことを狙いとしている。

(b) 社会的表示機構

聞き手へ向けた発話や話し手の心的態度を聞き手へ伝える上で、頷く、顔を向ける、身体を前後に傾ける等といった社会的表示は重要な意味を持つ。Talking-Ally は 4 つのサーボモータを内蔵することで、これらの表示を可能としている。

(c) 頭部・視線認識機構 (聞き手性のセンシング)

本提案システムでは聞き手の状態を把握するために、聞き手の「視線」を利用する。Goodwin によると、聞き手の視線には「今あなたの話を聞いていますよ」という表示を行い、話し手の発話を先に進める役割があると指摘する。このように、聞き手の視線は「聞き手性」の表示を担う一つであると同時に、聞き手の参加態度を推定するための重要な手掛かりになる。

聞き手の視線の認識システムとして視線・頭部追跡システム faceLAB5.0 (SeeingMachines 社) を利用する。これは図 2 のような仮想平面 (Virtual Planes for Eye-gaze) をソフトウェア上に作成し、その注目度合いにより発話方略を動的に選択し組織化する。(図 2 Dynamic Adaptation)

ここでは、Web から取得したニュース記事を読み上げるという発話生成システム(図2)を Talking-Ally に実装した。

3.2. 発話のデザイン

聞き手と相互行為的に調整しながら発話をデザインしていく上で、「他者からのアシストを引き出ししながら、結果として合目的な行為を実現する」という関係論的な行為形成の考え方を基にする[17]。

本発話生成システムは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が期待できる。

また、発話をデザインしていく上で発話を構築する方略が重要となる。本発話生成システムが考慮する発話方略として、以下のものがある。

(a) フレージング

私たちはある話題について発話を行うとき、一度にその全ての内容を伝えることはない。聞き手が聞き取りやすく理解しやすいように、その発話内容をある構成単位 (フレーズ) ごとに分割しつつ、発話を行うと考えられる。

これを実現するために、形態素解析エンジンである MeCab を用いて発話内容を構成単位ごとに分割 (フレージング) する。そして、発話を組織するための基本的な単位として利用する。

表1 ターン開始要素とモダリティ

Modality	Turn-initials	
1: 「ーね」	1: 「あーあ」	2: 「あー」
2: 「ーなんだって」	3: 「あのね」	4: 「あのさ」
3: 「ーらしいよ」	5: 「えーつと」	6: 「えつと」
	7: 「えつとー」	8: 「えつとね」
	8: 「ねーねー」	9: 「んつと」
	10: 「んつとね」	

表2 実験条件

条件	Modality/ Turn-initials	聞き手の視線への 配慮
A	無	無
B	有	無
C	無	有
D	有	有

(b) ターン開始要素 (turn-initials)

私たちは構成単位ごとに区切られた発話を行いながら、常に聞き手の会話への関心にも注意払う。そして、聞き手の視線が別の方向へ向いているとき、「えつと」、「あー」、「それでね」といったターン開始要素を発話する。2章の2.2節で議論したように、ターン開始要素には話し手が聞き手に「今から私が発話しますよ」ということを伝える効果がある。表1のようなターン開始要素を用いることで、聞き手の注目をTalking-Allyに引き戻したり、会話への参加態度を維持する狙いがある。

(c) 発話末のモダリティ (modality)

構成単位ごとに区切られた発話を行う場合、私たちは発話末に「〜ね」、「〜らしいよ」といったモダリティを付与する。中右の指摘[18]によれば、「文の意味内容は、つねに、命題(内容)とモダリティとの2つの領域に分けられる」としており、モダリティとは、「発話時という瞬間的現在の話者の心的態度を示すもの」と定義している。本研究では、この定義に従いモダリティを扱う。

発話を組織する際に表1のようなモダリティを発話末に付与し、聞き手のTalking-Allyに対する解釈をより強く引き出す。これにより、発話内容をTalking-Allyの内部で完結させるのではなく、一部を聞き手に委ね、発話の組織化への参加を促す狙いがある。

本発話システムでは、これらの方略に加えて言い直しや直接的な要求を用いて、聞き手の状態(=視線の変化)をリアルタイムに発話へ反映させながら、聞き手と相互行為的に発話の組織化を行う。

4. 評価実験によるデータの収録

4.1. 実験方法

本実験では、従来の発話生成システムと本提案シ

ステムにおいて、聞き手の振る舞いにどのような違いが現れるのかを比較・検証する。そのために表2に示す条件ごとに、発話生成システムを構築した。この4種類の発話生成システムを実装したTalking-Allyの発話を被験者に聞いてもらい、この一連の様子をビデオデータとして収録する。また、被験者の興味を散漫にさせるために、Talking-Allyの後方でビデオを上映した。

4.2. 被験者への教示

被験者は19歳から21歳の男性13名、女性1名の計14名である。被験者には「ロボットの会話を4パターン聞き、それぞれの会話が終了した後アンケートに答える」、「後方で流れているビデオの内容も記入する」という二つの教示を行った。また、実験の説明・教示をTalking-Allyに発話させることで、第1施行と第4施行でのTalking-Allyの音声への違和感や慣れを緩和させている。4種類の発話方法と発話内容は、カウンターバランスに配慮して被験者に提示した。

5. 収録データの分析

5.1. 動作例

ここで構築した発話生成システムを用いてユーザとインタラクションした際の動作例を次に示す。以下のトランスクリプト中の注釈の中で使用されているラベルの内容は表3の通りである。

なお、ここでの動作例においてはユーザ側からの発話は想定していない。被験者の視線の変化を手掛かりにロボットの発話をリアルタイムに変化させる動作をもって、ここでは「相互行為的な調整を含む発話」としている。ユーザ側の発話や身体の動き・ジェスチャなどを考慮に入れた相互行為的な調整は今後の課題とする。

5.1.1. インタラクション例1

インタラクション例1ではTalking-Allyが言い直しを行った場面を示している。まず、ユーザの視線

表3 トランスクリプト中のラベルとその内容

ラベル	内容
H	ユーザ
R	ロボット(Talking-Ally)
-forward	姿勢を前へ
-back	姿勢を後ろへ
-nod	うなずき
-human's gaze	人が向いている方へ視線を向ける
-gaze front	視線を正面に戻す
gaze in	Talking-Ally に視線を向けている
gaze out	Talking-Ally に視線を向けていない

が得られていない状態で、Talking-Allyが姿勢を前に倒し、話し始めるが、依然としてユーザの視線がTalking-Allyの方に向かなかつたため、再度同じ発話を行い、ユーザからの視線を得ようとしている。

以下に着目している場面のトランスクリプトを示す。

- 01 ((H: gaze out))
- 02 ((R: nod))
- 03 (0.8)
- 04 ((R: forward))
- 05 R: 沖縄警察のうるま署はね
- 06 ((H: gaze in))
- 07 ((H: gaze out))
- 08 ((R: back))
- 09 (1.0)
- 10 ((R: nod))
- ⇒ 11 R: 沖縄警察のうるま署はね
- 12 ((H: gaze in))
- 13 ((H: gaze out))
- 14 ((R: back))

5.1.2. インタクション例2

インタクション例2ではTalking-Allyが発話を躊躇した後で、言い淀みと、言い直しの両方を行った場面を示している。まず、ユーザの視線が得られていない状態で、Talking-Allyは姿勢を前に倒し、話し始めるが、依然としてユーザの視線が得られなかったため、それまでの発話を一時中断し、うなずきとともに「えーつとね」と発話し、ユーザの視線を得ようとしている。さらにユーザは、Talking-Ally以外の場所を多く見ていたため、「こっち見てよね」という直接的な要求をすることで、Talking-Allyはユーザの視線を獲得しようとしている。その後、先ほど躊躇して言い淀んだ発話を再開している。

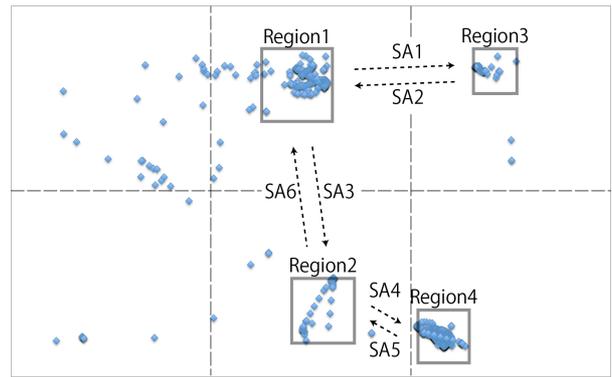


図3 faceLAB5.0による視線遷移データ

以下に着目している場面のトランスクリプトを示す。

- 31 ((H: gaze out))
- 32 ((R: forward))
- 33 (0.5)
- 34 R: 成人し
- 35 ((R: nod))
- 36 ((H: gaze in))
- 37 R: えーとね
- 38 ((R: nod))
- 39 ((H: gaze out))
- 40 (0.5)
- 41 ((R: forward))
- 42 (1.5)
- ⇒ 43 R: こっち見てよね
- 44 ((H: gaze in))
- 45 R: 成人式出席後にね
- 46 ((R: nod))

5.2. 視線遷移データ

条件Dの施行後半部分における、ある代表的な被験者の視線座標データをプロットしたものが図3である。Region1がTalking-Allyに相当する平面であり視線が集中している。また、その他にも特徴的なRegionがいくつか見られる。SAは視線遷移を表しており、番号順に視線遷移が見られた。

faceLAB5.0により得られた、こうした視線座標データを詳細に分析することで、本発話生成システムを実装したTalking-Allyが聞き手の視線の微細な動きや遷移に与える影響を調べることができる。また、従来の発話生成システムとの視線のばらつきについて、比較・検証することで本発話生成システムの有用性を明らかにできると考えている。



図4 ELANによるアノテーション例

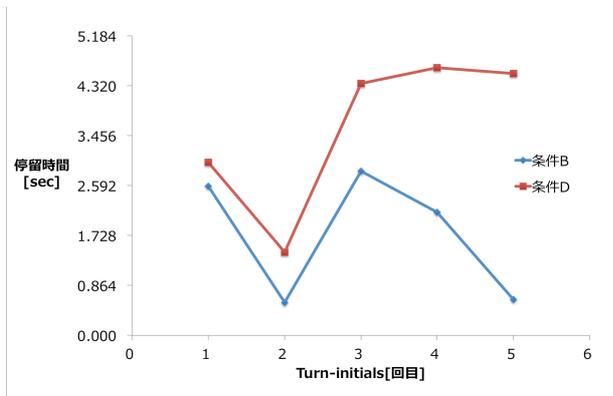


図5 代表的な被験者の停留時間遷移

5.3. 代表的な被験者の停留時間遷移例

収録したビデオデータを用いて、会話分析を行った。分析には ELAN (アノテーションツール) を利用して、ビデオデータで観察できるそれぞれの振る舞いに注釈を施した (図4)。視線遷移に関する注釈の時系列データから、ある代表的な被験者の Talking-Ally に対する視線停留時間 (Talking-Ally に注目している時間) の遷移を示したものが図5である。この図には、聞き手の視線に配慮しながら発話方略を用いる本発話生成システム (条件 D) と、聞き手に配慮せずに発話方略を機械的に用いるこれまでの発話生成システム (条件 B)、それぞれの視線停留時間の遷移が表示されている。このデータから次の傾向が見て取れた。

ターン開始要素3回目以降の遷移に注目すると、条件 B では Talking-Ally への興味を失ったかのように、停留時間が減少している。それに比べて条件 D では停留時間が3回目以降も、ほぼ同様に維持されていることから、3回目以降も Talking-Ally の話を聞こうという姿勢が維持されていると考えられる。こ

れらをまとめると、以下の可能性を示唆している。

- 聞き手の会話への参加態度の維持

条件 D では、聞き手の Talking-Ally に対する視線停留時間を維持していた。これから聞き手の参加態度が維持されているといえる。また、この現象は Talking-Ally の聞き手に配慮した発話方略を用いた発話に、聞き手の話題に対する興味を引き出すといった、聞き手を揺り動かすような「力」を備えることができる可能性を示唆する。

- 聞き手の構え (stance) が異なる可能性

条件 B と条件 D の停留時間の変化を比較すると、Talking-Ally に対する聞き手の構えに何らかの差異を感じる。これが「設計的な構え」と「志向的な構え」による違いだとすれば、詳細な分析を重ねることによって、志向的な構えを引き出すことが可能な発話生成システムのメカニズムを明らかにできる可能性がある。

6. まとめ

これまでのロボット研究では、ロボットの中身を作り込むことで、あらゆる機能をロボット一人で実現しようとする「個体能力主義」的な傾向が強かった。同様に、これまでの発話生成システムもその影響を色濃く受けている。本研究では、このような背景を持つ発話生成システムについて再検討し、「聞き手」と支えあいながら相互行為的に調整し発話の組織化を行う新たな発話生成システムを構想した。そして、そのプロトタイプとして「聞き手性」の一つである視線を「リソース」としながら、聞き手とともに発話を組織する発話生成システム「Talking-Ally」の構築を行った。

また、この発話生成システムを実装した Talking-Ally の評価を行う目的で、聞き手とともに発話を組織することで聞き手の振る舞いにどのような変化を与えるのか比較・検証を行った。また、これらの考察により得られた、聞き手の振る舞いの傾向と可能性について述べた。

今後は、これらの収録データの詳細な分析を進めるとともに、本発話生成システムに適応学習機構を実装することによって、聞き手の個性に適応した発話方略の生成過程の解析、会話における話し手と聞き手の距離感における分析や非流暢性の生成メカニズムを構成論的に検証していくことを考えている。

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。また、実験を行うにあたり、様々な被験者の協力を頂いた。ここに記して、感謝の意を表す。

参考文献

- [1] 浜田寿美男：『身体から表象へ』，ミネルバ書房（2002）。
- [2] 岡田美智男，三嶋博之，佐々木正人：『身体性とコンピュータ』，共立出版（2000）。
- [3] エドワード・S.リード著，細田直哉訳・佐々木正人監修：『アフォーダンスの心理学 生態心理学への道』，新曜社（2000）。
- [4] 鈴木健太郎，佐々木正人：行為の潜在的なユニット選択に働くタスク制約：日常タスクに観察されるマイクロスリップの分析，『認知科学』，Vol.8(2)，pp.121-138，（2001）。
- [5] 末崎裕康，古山宣洋，関根和生：アニメーションの物語説明における身振りにマイクロスリップは生じるか？—コミュニケーション場面におけるマイクロスリップに関する研究—，『日本生態心理学会 第3回大会 発表論文集』，（2010）。
- [6] 安田哲也，小林春美：マイクロスリップ”躊躇”の生起時間の変化と視線移動—日常作業時における熟達の影響—，『日本生態心理学会 第3回大会 発表論文集』，（2010）。
- [7] 中井陽子：言語・非言語行動によるターンの受け継ぎの表示，『早稲田大学日本語教育研究』，Vol.3，pp.23-39（2003）。
- [8] 山根智恵：日本語の談話におけるフィラー，『日本語研究業書』，Frontier series, Vol.15，くろしお出版（2002）。
- [9] 水上悦雄，山下耕二：対話におけるフィラーの発話権保持機能の検証，『認知科学』，Vol.14, No.4，pp.588-603（2007）。
- [10] Levelt, W.: *Speaking, From Intention to Articulation*, The MIT Press（1989）。
- [11] 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）。
- [12] Goodwin, C.: *Conversational Organization: Interaction between speakers and hearers*, Academic Press（1981）。
- [13] 中井陽子：日本語の会話における言語的／非言語的な参加態度の示し方—初対面の母語話者／非母語話者による4者間の会話の分析—，『早稲田大学日本語研究教育センター紀要』，Vol.19，pp.79-98（2006）。
- [14] 谷口忠大：『コミュニケーションするロボットは創れるか—記号創発システムへの構成論的アプローチ—』，communis, Vol.13，NTT出版（2010）。
- [15] Daniel C. Dennett（土屋俊 訳）：『心はどこにあるのか』，草想社（1997）。
- [16] B. J. Fogg（高良理，安藤知華 共訳）：『実験心理学が教える「人を動かすテクノロジー」』，日経 BP 社（2005）。
- [17] 岡田美智男：ゴミ箱ロボット—関係論的なロボットの目指すもの，『計測と制御』，Vol.51, No.8，pp.753-758（2012）。
- [18] 中右実：第4章 文副詞の比較，『日英語比較講座 文法』，Vol.2，pp.159-219，大修館書店（1980）。
- [19] 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）。
- [20] 岡田美智男：『口ごもるコンピュータ』，情報処理学会編 情報フロンティアシリーズ（9），共立出版（1995）。