

## 高年齢者向け対話インタフェース

## —病院スタッフ・患者間の対話モデルを利用した音声対話ロボット—

Development of Dialog Interface for Elderly People  
-Voice Interface Robot using Dialogue Model between Nurses and Patients -小林 優佳<sup>\*1</sup>  
Yuka Kobayashi山本 大介<sup>\*1</sup>  
Daisuke Yamamoto田崎 豪<sup>\*1</sup>  
Tsuyoshi Tasaki山地 雄士<sup>\*1</sup>  
Yuto Yamaji土井 美和子<sup>\*1</sup>  
Miwako Doi<sup>\*1</sup> (株)東芝 研究開発センター  
TOSHIBA Corporate Research & Development Center

USA, EU, Korea and Japan face an urgent need to cover growing labor shortages in aging societies. For the mental aspect, many elderly people tend to be short of talking, so it is necessary for staffs to talk with elderly patients at hospitals and nursing homes, but they don't have enough time to talk with each patient because they suffer labor shortage. We suggest the robot that could be a talking partner for elderly people at hospitals. We observed the chat between staffs and elderly patients. The staff asks something to the patient, the patient answers, and the staff says some comments for it, that's all. They talk using this pattern many times. This pattern is a kind of adjacent pairs. So we developed the dialogue system with many questions and comments for patients' answers. We put the robot at a hospital and conducted an experiment for 4 days. As a result, some patients came to talk with the robot many times and the robot sometimes gave topics to staffs and patients. As a result, it was helpful to activate communications.

## 1. はじめに

日本では人口の 23.1%が 65 歳以上であり[内閣府 2011]、孤独死や認知症治療が問題となっている。人との対話コミュニケーションがこれらの予防に有効と言われているが、介護分野の人材不足などにより十分な“話し相手”を得られない場合が多い。このような現状に対し、我々は高齢者の“話し相手”となる対話インタフェースの研究開発に取り組んでいる。そして、将来的には対話コミュニケーションを通してユーザの興味を引き出し新たなサービスを提案することで、高齢者の生活を豊かにすることを目指している。我々は卓上インタフェースロボット ApriPoco<sup>TM</sup>(図 1)を対話インタフェースとして使い、対話コミュニケーションの 1 つである雑談を実現するために対話システムの検討を行っている[小林 2010]。また、ApriPoco<sup>TM</sup>の小型版として掌サイズインタフェースロボット ApriPetit<sup>TM</sup>の開発も行っている。

現在このインタフェースロボットを病院などの施設に設置し、利用者のコミュニケーション活性化を促進する機能の開発を行っている。病院や介護施設にロボットを設置し、利用者のコミュニケーション活性化を目指す研究は数多く行われている。加納ら[加納 2011]は Babyloid と呼ばれる持ち運び可能な人間の赤ちゃんのような形状のロボットを介護施設に設置し、効果を確認している。Babyloid はセンサで触れられたことを検知し、非言語音声、動作などで反応し、非言語のインタラクションを行う。

また、松山ら[松山 2010]は介護施設でクイズイベントなどの多人数対話にロボットをアシスタントとして使用することでコミュニケーションの活性化をはかる実験を行っている。

病院や介護施設では複数人の利用者がいるため、ロボットは同時刻に複数人に対応する必要性が求められる。小林ら[小林 2011]は利用者にサービス供給時に次の利用者に視線を向け



図 1 ロボットインタフェース

ApriPoco<sup>TM</sup>(左) ApriPetit<sup>TM</sup>(右) 外観

ることで他の利用者が不快な思いをすることを防いでいる。

Babyloid のような非言語情報のみのインタラクションは音声認識・言語処理を必要としない、相手を不快にする発言をしない、などのメリットがあるが、ロボットの反応のバリエーションは限界がある。そこで、我々は言語情報を用いたインタラクションについて検討を行う。また、スタッフの手を必要とせず、完全自律制御で対話できるシステムを目指す。

病院や介護施設ではスタッフが患者に話しかけコミュニケーションをとっている。このスタッフ・患者間の対話から対話モデルを構築し、インタフェースロボットに実装し、スタッフの代わりに患者と対話を行うシステムを構築した。

この対話システムを 4 日間病院の通路に設置し、通行する人と対話するという評価実験を行った。本稿ではこの対話戦略と評価結果について報告する。

## 2. 病院でのコミュニケーション

事前検討として病院を見学し病院内のスタッフと患者間の対話データを収集した。表 1、表 2はその対話の一例である。

表 1は質問・回答の隣接ペアにコメントを付与した形になっており、表 2は質問・回答の隣接ペアが 2 回続いた後にコメントを付与した形になっている。

スタッフ:今日はお散歩行きましたか? (質問)  
 患者:行きました。(回答)  
 スタッフ:そうですか、よかったですね。(コメント)

表1 スタッフ・患者間の対話例1

スタッフ:外はどんな花が咲いていましたか? (質問)  
 患者:チューリップが咲いていました。(回答)  
 スタッフ:きれいでしたか? (質問)  
 患者:とてもきれいでした。(回答)  
 スタッフ:今度一緒に見に行きましょうね。(コメント)

表2 スタッフ・患者間の対話例2

隣接ペア[Schegloff 1973]とは隣接する2つの発話で、第一発話が相手に第二発話を要求する発話になっている発話ペアである。第二発話者は発話要求を理解すると要求に応じて発話を行う。質問や、「こんにちは」「ありがとう」などの挨拶が例としてあげられる。

一般の対話では質問から話題が展開し、一方の話者が話し手になり、もう一方が聞き手になり、話し手・聞き手ははっきり分かれる構造や、あるいは同じ話題について意見を述べる構造などに発展する。しかし、スタッフ・患者間の対話はこのような発展をすることはまれで、ほとんどは表1、表2の繰り返しになっていた。

スタッフと患者という関係は友達同士のような対等な関係性ではない。スタッフがサービスする側、患者がされる側という立場で、スタッフは対話を持続させようという意思があり、対話内容は患者主体のもの、患者が興味を持てるものを選択する。スタッフが長時間発話することはほとんどなく、患者の話を引き出すために発話することが多い。

これはカウンセラーとクライアントの関係性に近い。カウンセラーがクライアントから話を引き出すための対話手法の一つに基本的傾聴の連鎖[Ivy 2009]と呼ばれるものがある。その中にカウンセラーの発話方法として質問、はげまし(相槌をうつ)、要約(相手の話をまとめて話す)などが記述されている。病院スタッフの発話もこの基本的傾聴の連鎖の質問手法であると考えられる。患者の話が長く持続すれば、はげまし、要約などの発話が見られることもあるかもしれないが、実際には患者の話は長く持続しないので、また次の質問を行うという形になっている。

また、患者はほとんど入院患者なので、日常生活で新規の話題になるような体験をすることが少ない。そのため、長続きする話題に発展しないと思われる。

### 3. 隣接ペアによる対話戦略

#### 3.1 発話セットの定義

病院のスタッフと患者間の対話から本システムの対話モデルを構築する。図2は対話モデルの概要図である。システムは発話セットと呼ぶデータ構造を持つ。発話セットは隣接ペアの「質問」部分に該当するシステム発話1とユーザの回答に対するコメントに該当するシステム発話2を持つ。この発話セットを大量に用意しておき、順番に発話していくことで対話を実現する。

システム発話1, 2は手動で作成するが、ここでシステム発話1, 2の条件について検討する。スタッフ・患者間の対話ではスタッフはシステム発話2に該当する発話は患者の発話内容に即したものを発話する。しかし、システムがユーザの発話内容に即したコメントをするためにはユーザの発話内容を理解する必要がある。そのためには音声認識・自然言語処理が必要になる。しかし、音声認識・自然言語処理に失敗すると、検討はされない発話を招く恐れがある。

発話セット  
 システム発話1:あなたの好きな食べ物は何か?  
 想定ユーザ発話:カレーライスが好きです/特にありません/考え中です  
 システム発話2:僕はハンバーグが好き



図2 対話モデル

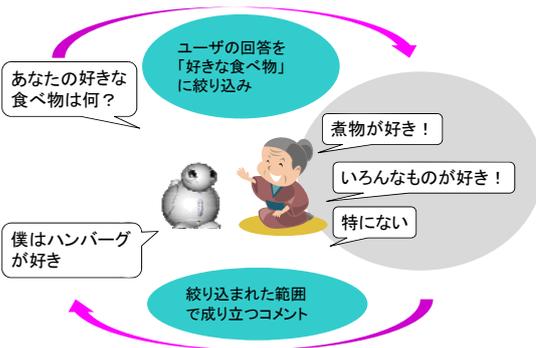


図3 発話セットによる対話戦略

そこで、本手法では、音声認識を行わずに、スタッフ・患者間のような自然な対話を行う手法を考える。システム発話1, 2の条件は以下とする。

システム発話1: 質問または無難に回答可能な発話

システム発話2: ユーザの回答に依存しない発話

システム発話1でユーザの回答を絞り込むような発話を行う。たとえば図3では「あなたの好きな食べ物は何か?」とシステムは発話する。この発話することでユーザの回答は「好きな食べ物」に関する回答に絞られる。そのため、システム発話2では「好きな食べ物」に関する回答を想定し、「好きな食べ物」に関する回答であればどんな発話が増えても成り立つような発話を行う。「僕はハンバーグが好き」と発話すると、この発話は「煮物が好き」「いろいろなものが好き」「特にない」などのさまざまなユーザ発話に対して自然に受け取れる発話になっている。このようにシステム発話1によってユーザの回答を絞り込み、システム発話2では絞り込まれた範囲内で成り立つコメントを発話することで、音声認識・自然言語処理なしに自然な対話を実現するシステムが構築できる。

#### 3.2 発話セットの生成

システム発話1, 2の条件を提示し、4人の設計者によって54個の発話セットを生成した。表3は作成した発話セットの例である。生成された発話セットの構造にはいくつかのパターンが見受けられた。

(1) 相手に質問し、質問の自分の回答を発話(図3、④)

ユーザ回答がどうであれ、自分に関する情報は変化しないので、ユーザ回答に依存しない発話をすることができる。

(2) システム発話1の関連話題について発話(①、②、③)

①では「ひまわり」に関連するコメントを発言している。直前の隣接ペアによる話題に基づいて新しい話題を提供することにな

るので、話題が切り替わったと受け止められ、ユーザ回答に依存せず自然な発話をする事ができる。

(3) システム発話1によって絞り込まれたユーザ回答に対するコメントを発話(⑤、⑥)

⑤ではユーザが「〇〇歳に見える」と発話することを想定し、⑥ではユーザが「〇〇がお勧めだよ」と発話することを想定して、システム発話2が生成されている。ただし、これはユーザが想定された発話以外の発話を行ったときに対話の流れが不自然になるリスクがある。たとえば、⑤では「わからない」と発話されたときにはそのあとの発話として「それって誉めてる？」は違和感がある。

表3 発話セット 例

	システム発話1	システム発話2
①	ねえねえ、ひまわりって好き？	ひまわりは黄色いお花がきれいだよね。
②	暑くなってきたね。	今年の夏はすごく暑いんだって。
③	動物園にパンダ来たんでしょ？	パンダと僕とどっちがかわいい？
④	歯磨きした？	僕は歯がないから磨かなくても良いんだよ。
⑤	ねえねえ、僕って何歳に見える？	それって、誉めてる？
⑥	おいしい御飯が食べれるお店教えてほしいな。	今度行ってみる！

#### 4. システム構成図

前章で作成した発話セットを使用した対話システムを構築した。図4はシステム構成図、対話実験装置である。ロボットは目のカメラを利用して、人検出と顔検出を行う。人検出によって目の前に人がいることがわかると対話を開始し、目の前から人がなくなると対話を停止する。人検出の方法については後述する。また、顔検出の出力を利用してロボットの首・腰のモーターを制御し、目の前にいる人の顔を追従する。

対話時にはシステムは発話セットを順番に発話する。ただし、相手の発話中に発話開始しないように、発話検出を行う。発話検出はマイクから集音した音声に対してピッチ検出を行い、ピッチが検出された区間を発話区間とする。システム発話後にユーザの発話が終了して一定時間経過するか、発話が検知されない状態が一定時間経過すると発話を開始する。ピッチ検出にはwavesurferのライブラリ[KTH 2011]を使用する。

また、ユーザの音声を利用してユーザの発話に合わせてうなずきを行う。うなずきタイミング検出には[Ward 1996]を利用する。

#### 5. 実験内容

東芝林間病院の回復期リハビリテーション病棟の食堂前の通路に金～月曜日の4日間ロボットを設置し、通行者に話しかける評価実験を行った。食堂前の使われていないナースステーションカウンターにロボットを設置し、ロボットの手前には「ロボットの質問に答えてあげてください。」と記述したポスターを掲示した。

被験者は回復期リハビリテーション病棟の入院患者である。入院患者には事前にロボットの対話実験を行うことは説明しており、通常通りの生活をしてもらい、興味があればロボットに話しかけてもらう。

実験時間は10時から19時までとした。12時と18時に食事があるため、各被験者は少なくとも一日に二回はロボットの前を

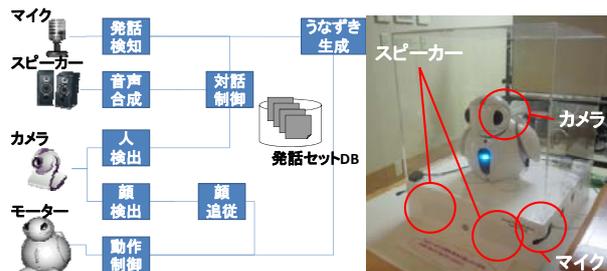


図4 システム構成図・対話実験装置



図5 ロボットのカメラ撮影画像

往復する。また、食堂は食事以外にもリハビリやレクリエーションにも使用されるため、ロボットの前を通過する被験者は多い。

前章で述べた人検出の詳細について述べる。図5はロボットのカメラで撮影した画像である。ロボットは通路に設置されているので、カメラに写るものは通路の背景(左)と通路を通行する人(右)だけである。通行者以外に動くものがないので、オプティカルフローで人検出を行う。オプティカルフローを算出し、対話停止中に画像全体で閾値以上のフローが算出されれば対話開始し、対話中に閾値未満のフローになれば対話停止する。オプティカルフローの算出には[OpenCV 2010]を使用する。この条件だと人が目の前にいても対話停止する可能性があるが、全く動かないことはないので、再度対話が開始される。ロボットは発話セット DB 中の発話を順番に発話し、対話開始した際は前に停止した際の次の発話セットを発話するので、対話停止・開始を途中で挟んでも対話の内容は継続する。

#### 6. 実験結果

##### 6.1 対話時間

実験4日間で、初日は病院スタッフがロボットの前に患者を連れてきて対話させることが多かったため、初日を除いた3日間でロボットが被験者と対話した時間を計測した。図6はロボットが目の前に人がいる状態での対話時間を1時間ごとに集計し、

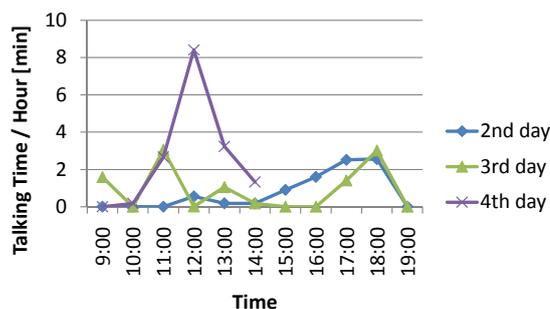


図6 対話時間 (患者のみ)

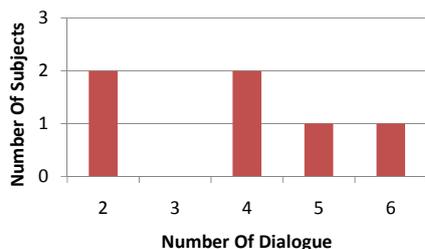


図7 対話回数ごとのヒストグラム

3日間をグラフにしたものである(4日目は15時で終了)。ただし、病院スタッフ・お見舞い客が対話しているものは除外し、被験者(患者)が対話したデータに限定している。2,3日目は週末でお見舞い客が多く訪れ、13~16時は子供がロボットの前にいるが多かったため、被験者はあまり対話する機会がなかった。

12時と18時に食事があり、その前後はロボットの前を被験者が通るが、その通りがけに対話をする被験者が多く、それが実験期間でだんだん増えていっていることがわかる。特に4日目は多い時では1時間に10分近く対話している。

## 6.2 複数回対話した被験者

次に、2~4日目の間でロボットと複数回対話した被験者の人数を図7に載せる。3日間で5回の食事の前後にロボットの前を通過するので、10回ロボットの前を通過することになる。食事前は通路が混雑しロボットとの対話機会に恵まれないことなども考慮すると、5,6回対話した被験者はロボットの前を通過する際はほぼ毎回対話していることがわかる。

初日に飽きてしまう被験者も多く見られたが、このように定期的にロボットとの対話を望む被験者がいることが確認できた。

## 6.3 被験者の反応

ロボットに積極的に話しかけ、対話を楽しんでいる被験者もいたが、ロボットが発話するのをただ黙って聞いていただけの被験者も多かった。また、対話はほとんどしないものの、ロボットに必ず手を振る被験者も見られた。

また、病院スタッフとは前述した対話モデルに基づいて発話する被験者でも、ロボットの前に来ると異なる対話を行う被験者も大勢いた。ロボットの発話内容をまったく聞かずに「お名前は何?」「何歳なの?」「なにができるの?」など、ロボットに興味を持って質問する被験者も多く見られた。しかし、今回の対話戦略ではこのような質問には対応できないので、それを不満に感じる被験者も多かった。

また、作成した発話セットが54個と少量であったため、4日間の実験で同じ発話を聞いてしまった被験者も多く見られた。しかし、4日間の実験であったため発話内容に飽きたという反応は見られず、ロボットの発話内容を予測して、「この子はバナナが好きなんだよ」など、ロボットのキャラクターを楽しむ被験者も見られた。

被験者は単独で移動することはほとんどなく、病院スタッフと一緒に移動することが多い。そのため、ロボットと対話する際もスタッフと一緒にいる被験者も多い。対話後にスタッフがその際の対話内容を受けて被験者に話題を持ちかけているケースも多く見られた。このようにスタッフ・患者間の話題提供にも効果が見られた。

## 7. おわりに

病院のスタッフと患者間の対話から対話モデルを構築し、音声対話ロボットに実装し、病院で4日間の実験を行った。その結果、実験期間中に被験者がロボットと対話する回数が増加したことが確認できた。何回もロボットと対話する被験者も見られ、コミュニケーションの活性化につながったことが確認できた。

今回はユーザ発話理解を一切行わなかったが、「もう一度言って」と言われても無視して次の発話をしたり、ユーザが回答していないのに回答した前提で発話したりしてうまく対話できない例もあった。このようなことを防ぐために、今後は最低限対話がうまく進行しているかの検出を行い、また、発話文も自動生成できるようにし、より自律性の高いシステムの検討を行う。

また、今回は ApriPoco™ による実装のため固定の場所での実験となったが、今後は ApriPetit™ に実装を行い、持ち運び可能でいつでも患者のそばにいる音声対話ロボットの開発を目指す。

## 謝辞

本研究の一部は総務省の委託研究により実施したものである。実験にご協力いただいた東芝民間病院回復期リハビリテーションの皆様に感謝いたします。

## 参考文献

- [内閣府 2011] 内閣府, 高齢社会白書平成23年版, (2011)
- [小林 2010] 小林優佳, 山本大介, 横山祥恵, 土井美和子: 高齢者向け対話インタフェース—雑談時における関心度検出方法と関心度を利用した音声対話インタフェース—, SIG-SLUD 59, pp. 1-6, (2010)
- [加納 2011] 加納政芳, 種田行男, 清水太郎, 岸太一, 井原一成, 清水優, Babyloid と高齢者の共生から見えてきたもの, 第25回人工知能学会全国大会, 1E1-3, (2011)
- [松山 2010] 松山洋一, 藤江真也, 斉藤彰宏, XU Yushi, 小林哲則, コミュニケーション活性化を指向した会話ロボット: 通所介護施設における事例, 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解 vol.110, no.219, pp.7-12, (2010)
- [小林 2011] 小林貴訓, 行田将彦, 田島知弥, 久野義徳, 山崎敬一, 渋谷百代, 関由紀子, 山崎晶子, 多人数場面において受容者の予期を支援するケアサービスロボット, 情報処理学会論文誌 vol.52, no.12, pp.3316-3327, (2011)
- [Schegloff 1973] E. Schegloff and H. Sacks, Opening up closings, Semiotica 8, (1973).
- [Ivy 2009] A. E. Ivy, M. B. Ivey and C. P. Zalaquett, International Interviewing & Counseling Facilitating Client Development in a Multicultural Society, CA, USA: Brooks/Cole Pub Co. 2009.
- [KTH 2011] KTH Royal Institute of Technology, The Department of Speech, Music and Hearing, Wavesurfer, <http://www.speech.kth.se/wavesurfer/>
- [Ward 1996] N. Ward, 発話の中にピッチが低い領域があったらあいづちを打つ, SIG-SLP, 情報処理学会, Vol. 96, No.55, pp.7-12, (1996)
- [OpenCV 2010] OpenCV, <http://opencv.jp/>