

認知症患者のための語りかけエージェント —プロトタイプシステムによる実証研究—

Conversational Agents for Dementia: A Field Study for a Prototype System

中野 有紀子*¹
Yukiko Nakano

比企野 純太*¹
Junta Hikino

安田 清*²
Kiyoshi Yasuda

*¹ 成蹊大学 理工学部情報科学科
Dept. of Computer and Information Science, Seikei University

*² 千葉労災病院/京都工芸繊維大学
Chiba Rosai Hospital / Kyoto Institute of Technology

With the goal of developing conversational agents that can serve as a conversational partner for Dementia, first, this study conducted a data collection experiment in WOZ setting. In a preliminary analysis, we found that our animated agent was accepted by the subjects. Then, we developed an autonomous conversational agent that can generate positive feedback, such as head nod and verbal acknowledgement based on the acoustic information in user's speech: pitch and intensity. Our system also judges whether the user finished speaking or not, and avoids interrupting user's utterance. Finally, we evaluated our system and found that the implemented conversational agent was accepted by subjects with Dementia. However, it was suggested that the social signals were not properly perceived because the amount of subjects' speech was not different depending on whether social cues were displayed or not.

1. はじめに

高齢者の人口が急速に増え、高齢化社会における介護者の不足は重大な問題であり、高齢者用施設、病院、また家庭等様々な場面において、高齢者への支援が必要とされている。特に、認知症患者は高齢であることに加え、記憶障害のために、薬の飲み忘れ、同じことを何度も尋ねる、時として精神的に不安定な状態になるといった状況が起り、介護者の負担が増える原因となっている。

一方、情報技術やロボット技術による高齢者の支援が試みられており、高齢者の自立支援、見守り、高齢者間のコミュニティづくり等を目的とした、社会的ロボットの研究・開発がすすめられている[6, 7]。このようなコミュニケーションロボットの研究では、ロボットが高齢者のコミュニケーション相手となるために必要な機能とは何かを探求し、それをロボットのインタラクション機能の設計指針とすることが重要である。

そこで本研究では、認知症患者のコミュニケーション相手となることで、介護者の負担を軽減させ、患者に精神的安定を与えることができる擬人化エージェント(ロボットやアニメーションのキャラクタ)の実現を目指し、認知症患者のための語りかけエージェントのプロトタイプを作成した。また、認知症患者とアニメーションエージェントとのインタラクションコーパスを収集し、今後のシステム改良に向けた分析結果についても報告する。

2. コミュニケーション相手として受容されるための要件

認知症患者のための語りかけエージェントを実装するためには、まず、ユーザである認知症患者がエージェントを話し相手として受け入れることが必要条件となる。この点において、先行研究では、擬人化エージェントが人間と同様の社会的シグナルを表現できることが重要であり[1]、それによりエージェントへの信頼性や親密性が向上することが示されている[2]。また、[3]では、これらの研究に基づき、擬人化エージェントに求められる社会

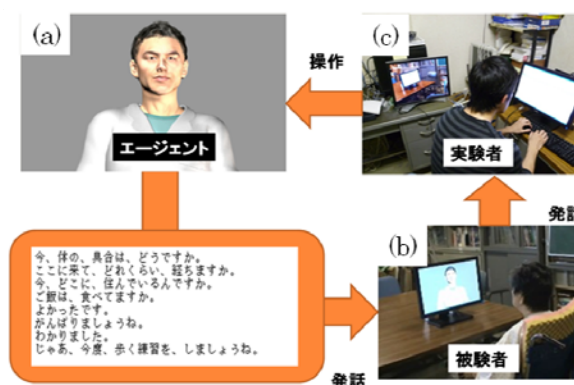


図 1 : WOZ 実験

的行動として、以下の項目を提案している。

1. 相手の話に注目していることを示す。(例)相手を見る、頷く
2. コミュニケーションに対して肯定的であることを示す。(例)笑顔で話す
3. 個人的なことを覚えていることを示す。(例)相手の名前を呼ぶ
4. 表現力をもって話す。(例)表情を使う
5. 間違いを認める

しかし、これらは健常者を対象とした研究において得られた知見であり、これらが認知症患者のためのエージェントのデザインにおいても有用であるのか否かを検証する必要がある。また、認知症患者は擬人化エージェントという人工物に対してどのような反応を示すのかについても調査すべきであると考えた。そこで、次節では、Wizard-of-Oz (WOZ) での語りかけエージェントとの会話実験について、また 4 節では、上記 1 の点に焦点を当てた、語りかけエージェントの実装について述べ、5 節ではこれを用いた評価実験の結果について報告する。

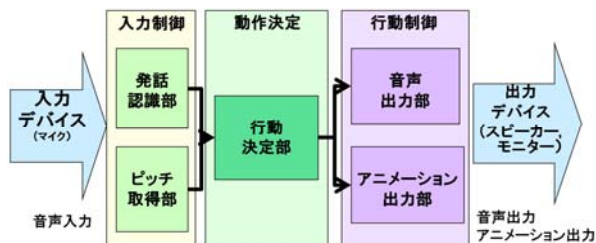


図 2: 語りかけエージェントのシステム構成

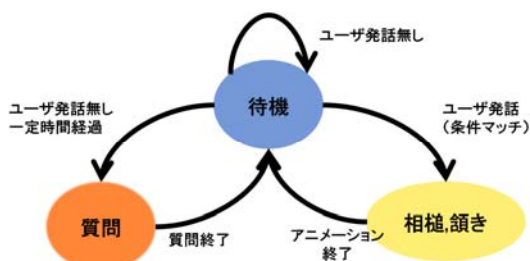


図 3: エージェント状態遷移図

3. WOZ 実験

本節では、認知症患者は擬人化エージェントに対してどのような反応を示すのか、コミュニケーション相手として受容する可能性があるのか否かを検証するために、実験者がエージェントの応答を制御する WOZ 実験により、データ収集を行った。

3.1 アニメーションエージェントの作成

本研究では、擬人化エージェントとして、ロボットではなくアニメーションのエージェントを作成した。キャラクタの 3D モデル作成には Poser7 を用い、今回の実験は病院内で実施したため、病院スタッフを想定し、白衣を着た日本人男性のキャラクタを使用した。実験に使用したキャラクタのスナップショットを図 1(a)に示す。エージェントは大きな動作は行わず、発話中の口の動き、頷き動作、またアイドリング動作として微細な体の揺れを表現することとどめた。これらのアニメーションは Poser7 で作成されたアニメーションから画像ファイルを書き出し、これらをキーフレームとして Adobe Flash ver.10.0 により、web ブラウザ上で連続描画することにより実装した。

3.2 実験の実施

被験者はエージェントが映っている 20 インチディスプレイの前に座り(図 1(b))、実験の様子を背後からビデオカメラで撮影した。その映像が別室にいる実験者のモニタに表示され(図 1(c))、それを見ながら実験者がエージェントを操作することにより、エージェントの発話や相槌等の動作が実行された。エージェントの発話はディスプレイ横のスピーカーから出力されたが、聴力に不安のある被験者にはイヤホンを使用していただいた。また、被験者に尋ねる質問は、千葉労災病院の認知症患者の介護にあたっているボランティアの方の認知症患者との対話の書き起こしから抽出した。図 1 に質問の例を示す。

実験の被験者は、千葉労災病院の整形外科に入院中の患者 2 名と物忘れ外来に通院中の認知症患者 3 名の計 5 名であ

る。全員女性で平均年齢は 79.8 歳で平均 MMSE は 12.2 である。

実験は、エージェントからの全質問が終了した時点で終了とした。実験終了後に被験者に対して簡単なインタビュー形式のアンケートを行った。対象者とその家族には実験の目的と内容を説明し、実験協力への同意をいただいた。

3.3 予備実験からの知見

WOZ による予備実験では全員の患者において、用意した質問全てが実行され、会話は問題なく遂行された。質問に対するレスポンスも早く、時折笑顔を見せる被験者もいた。また、実験後のインタビューや観察から以下のような知見が得られた。

- 全員の患者において、エージェントとの会話が成り立つことが確認された。このことより、エージェントによる相槌や頷きといった動作は健常者の場合と同様、有用な社会的シグナルとなる可能性がある。
- 会話エージェントに対して被験者から質問する場面は 1 度も見られなかった。従って、エージェント側からユーザーが興味を引くような質問を適切なタイミングで行い、相手の返答をうまく引き出す機能が必要であると考えられる。
- エージェントの発話タイミングが悪く、被験者の発話とオーバーラップした場面が複数回見られた。従って、被験者の発話終了を待って、次の問いかけを行うことが重要であると考えられる。

以上の観察結果より、認知症患者に対してアニメーションエージェントが語りかけ対話することにより十分コミュニケーションが成り立つことが分かった。また、全ての被験者が落ち着いた状態で実験を完了できたことから、語りかけエージェントが有効であることが示唆された。

4. 語りかけエージェントの実装

予備実験より、認知症患者と WOZ エージェントとのコミュニケーションは予想以上にスムーズであることが確認されたため、予備実験で得られた知見に基づき、エージェントの自動化に取り組んだ。本研究では、最も基本的な社会的シグナルとして、相手(患者)の発話中の相槌(「ええ」「はい」)、無言での頷きを表出する機能に加え、相手の発話終了を待って応答を返すために、発話終了認識機能を実装した。システムの構成を図 2 に示す。

入力制御モジュール:発話認識部ではマイクから入力された音声の音量から、発話入力の有無を判定する。声の大きさは人によって異なるので、判定の閾値は適宜調整することとした。次に、ピッチ取得部では、発話入力であると判定された音声に対して、ピッチを算出する。これらの値は動作決定モジュールに出力される。

動作決定モジュール:行動決定部では、入力制御モジュールから得られた音声情報を用いて、エージェントの状態の更新と行動の決定を行う。図 3 にエージェントの状態遷移図を示す。エージェントは質問を実行すると待機状態に入り、ユーザの発話が発見されると、入力制御モジュールから得られたピッチの情報をフィードバック生成ルール(表 1)に適用し、相槌と頷きを決定する。フィードバック生成ルールは[4]と[5]を参考にして作成した。頷きは最新の 16 個のピッチ情報から決定されるため、ユーザの発話途中にも実行されるが、相槌は発話終了ごとに決定される。相槌や頷きが実施されると、また待機状態に戻る。エージェントの質問後、一定以上の時間が経過してもユーザからの応答がない場合には、次の質問に移行する。

行動制御モジュール:行動制御モジュールは、キャラクターのアニメーションの実行と音声ファイルの再生を行うが、これには WoZ 実験のために実装した機構をそのまま利用した。

表 1:フィードバック生成ルール

条件	エージェントの行動
現在のピッチ n が $n-1$ から $n-16$ までの平均ピッチより 5%低下	頷き
発話終了後 2 秒以内の無音があり再度発話が検出される	相槌(ええ)
発話終了後 2 秒以上の無音 or 発話終了直前のピッチ下り幅が 2%以内 or 発話終了直前のピッチが発話平均ピッチより 10%以上上昇	相槌(はい)

5. 評価実験

先の WoZ 実験の結果を考慮して作成した会話エージェントシステムを用いて認知症患者との対話実験を行った。本実験により、認知症患者とエージェントの対話において、エージェントによるフィードバックが有効であるかを分析するためのデータ収集を行う。

5.1 実験条件

ユーザの発話タイミングを考慮した語りかけの有効性、および相槌や頷きによるフィードバックの有効性を検証するために、以下の 3 種類の会話エージェントを用意し、各被験者に全てのエージェントと対話を行ってもらった。

- (1) 協調的エージェント:相手の発話音声が続いている場合は、相槌や頷きを返すことにより、相手の発話に反応し、6 秒以上発話音声がない場合に次の質問に移る。
- (2) 非協調的エージェント:相手の発話に関わらず、10 秒間隔で質問を発話する。相手の発話中には相槌や頷きは行わない。
- (3) モノログエージェント:相手の発話の有無に関わらず、10 秒間隔で一方向的に発話をする。

この 3 条件により、(1)、(2) 間では相槌の有無と質問のタイミングによる差異が、(2)、(3) 間では語りかけ方(一方向的に話す vs. 質問する)の違いによる差異が現れると予想した。

5.2 実験手続き

被験者は、千葉労災病院の物忘れ外来に通院中の認知症患者男性 6 名、女性 7 名の合計 13 名、平均年齢は 77.6 歳である。MMSE の値は 11~29、平均値 22.6 で比較的軽度の方が多かった。

実験時には、「この(画面を指差し)エージェントが話しかけてくるので、会話してみてください」と簡単な教示を伝えた。エージェントによる質問群は 2 種類用意し、条件(1) (2)にランダムに割り当てた。また、条件の実施順序も偏りのないようにした。1 セッションは 7 分とし、全ての質問が終わっていない場合でも 7 分を経過するとエージェントが終わりのあいさつをし、会話を終了するように設定した。実験終了後、口頭でのインタビュー形式でのアンケートを行った。実験協力者とその家族には実験の目的と内容を説明し、了承を得た。

被験者の正面と横からビデオデータ撮影すると同時に指向性の高いピンマイクにより音声データを収録した。

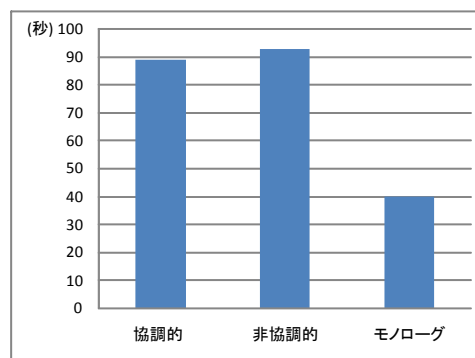


図 4: 発話量

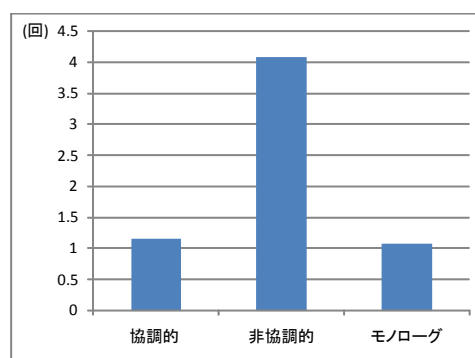


図 5: オーバーラップ回数

5.3 分析結果

収録したビデオデータを課題ごとに分割し、音声分析ソフト Praat を用いて発話区間のラベリングを行った。さらにこのアノテーション結果をビデオ分析ソフト Anvil に読み込み分析用データを作成した。

(1) **発話量:**被験者のセッションあたりの平均発話量(秒)のグラフを図 4 に示す。協調的エージェントに対する被験者の平均発話量は 88.9 秒、非協調的エージェントに対する被験者の平均発話量は 92.8 秒であり、両エージェント間に大きな差は見られなかった。一方、モノログエージェントに対して被験者が語りかけることは少なく、被験者の平均発話量は 40.1 秒とエージェントが質問をする条件と比べて半分以下であった。

(2) **発話のオーバーラップ回数:**被験者の発話とエージェントの発話とのオーバーラップ回数のセッションごとの平均を図 5 に示す。協調的エージェントが平均 1.15 回、非協調的エージェントが平均 4.08 回、モノログエージェントが平均 1.08 回となった。また、協調的エージェントと非協調的エージェントの平均オーバーラップ回数の差について t 検定を行ったところ、統計的に有意な差がみられた($t = -3.37061$, $p < 0.01$)。

5.4 考察

協調的エージェントと非協調的エージェントの間で、被験者の平均発話量に差が見られなかったことから、エージェントによる相槌や頷きの社会的シグナルの表出が、必ずしも認知症患者の発話量増加につながるわけではなかった。実際、被験者がこの 2 つのシステムの違いに気づいていない様子も見られた。そこで、病院スタッフの方にも本システムを使っていたところ、協調的エージェントと非協調的エージェントの違いにはすぐ気づいたことから、認知症患者と健常者とはエージェント

からの社会的シグナルの認知に違いがある可能性が示唆された。

一方、非協調的エージェント条件では、協調的エージェント条件と比べて平均オーバーラップ回数が有意に多かったことから、被験者自身が意識したかどうかは不明であるが、協調的エージェントの方が会話をしやすかったのではないかと推測される。

モノログエージェントに対しては、被験者の反応が大きく異なっていた。このことから、語りかけエージェントは、一方的に話すのではなく、ユーザに問いかける形態でのインタラクションを行う方が望ましいと言える。また、エージェントは単純に質問するのではなく、「私はみかんが好きですが、あなたはどんな食べ物が好きですか」のように、エージェントが自分自身の話も交えながら質問する方がより被験者からの発話を引き出しやすいことがわかった。エージェントが自分自身のことを先に言うことにより、ユーザに対して回答例を提示することになり、ユーザはよりこたえやすくなると考えられる。

また、実験中の被験者の反応を観察すると、興味のある話題とそうでない話題とで、発話量が大きく異なることがわかった。質問の内容についても、被験者の興味に合った話題を選択できるようにする必要があると考える。

6. まとめ

本研究では、認知症患者の精神安定、介護者の負担軽減を目指して、認知症患者支援用の会話エージェントを実装した。今後は、評価実験で収集したデータの分析をさらに進め、発話量だけでなく、被験者の表情等についても評価指標に加えていく予定である。また、分析結果に基づき、エージェントの動作決定ルールを改良する必要がある。また、相手の興味に合わせてエージェントの質問内容を変更する等の質問選択における工夫も進めていく予定である。

謝辞:実験の実施にご協力いただくとともに、貴重なご意見をくださった千葉労災病院リハビリテーション科のスタッフの皆様、実験に参加して下さった患者様に心より感謝いたします。

参考文献

1. Breazeal, C., *Towards sociable robots*. Robotics and Autonomous Systems, 2003. **43**(3-4): p. 167-175.
2. Bickmore, T., and Picard R. W. *Towards Caring Machines*. in *CHI 2004*. 2004.
3. Heerink, M., et al., *Studying the acceptance of a robotic agent by elderly users*. International Journal of ARM, 2006. **7**(3).
4. Gratch, J., et al., *Virtual Rapport*, in *6th International Conference on Intelligent Virtual Agents*. 2006, Springer: Marina del Rey, CA.
5. Tsukahara, W. and N. Ward. *Responding to Subtle, Fleeting Changes in the User's Internal State*. in *CHI 2001*. 2001: ACM.
6. Wada, K., Shibata, T., Saito, T. and Tanie, K., *Effects of Robot Assisted Activity to Elderly People who Stay at a Health Service Facility for the Aged*, Proceedings of the 2003 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, Las Vegas, Nevada, October 2003.

7. 米澤朋子他, 高齢者同士の TV 会話におけるユーザの参加状態に応じた介在型ロボット, HAI シンポジウム 2010.