

発話中の非言語情報に基づく認知症高齢者の状態把握

Assessing the Health and Mental Status of the Elderly with Dementia based on Nonverbal Information in Spoken Utterances

野中 裕子*¹ 高橋 明秀*¹ 林 佑樹*² 中野 有紀子*²
Yuko Nonaka Akihide Takahashi Yuki Hayashi Yukiko Nakano

*¹ 成蹊大学大学院理工学研究科
Graduate School of Science and Technology, Seikei University

*² 成蹊大学理工学部
Faculty of Science and Technology, Seikei University

Assessing the elderly's health and mental status and reporting it to caregivers is one of the key issues for supporting elderly people with dementia. Aiming at contributing to this issue, this study proposes a mechanism that assesses elderly's responsiveness in conversation with a virtual agent. Our assessing method uses prosodic information in their speech and by applying clustering technique, we create a model for judging whether each utterance is high or low in its responsiveness. We also develop a report system that automatically generates an assessment digest for the conversation, and a user interface for playing a video for a specific part of the conversation.

1. はじめに

少子高齢化が急速に進む中、高齢者支援の必要性が高まっている。特に、認知症高齢者は記憶障害のために介護者の負担が大きく、このような認知的障害を持つ高齢者の支援に注目が集まっている。例えば、高齢者が安全に日常生活を送っているのかを確認したり、生活を送るにあたって必要な動作を支援、また高齢者の認知状態を評価するといった支援が必要と考えられている[Pollack 2005]。そこで、患者の健康状態や精神状態を把握し、様子に変化があれば介護者に連絡する見守りサービスを実現するための技術として、本研究では、会話エージェントとの会話を通して、対象者の状態把握を行う技術の確立を目指す。

具体的には、エージェントからの問いかけに対する「反応性」に着目する。問いかけに対してテンポよく、リズムよく、また積極的に応答している様子を反応性が良いとし、発話音声から反応性が良いか悪いかを自動的に判定する技術を開発する。

さらに、反応性の判定手法の応用として、会話セッションのデータを自動的に解析し、反応性の評価結果をまとめたレポートを自動生成するレポート生成システムを実装する。本システムを遠隔地の医師やご家族に利用していただくことにより、認知症高齢者の様子を手短かに知らせることが可能になる。

2. 語りかけエージェント

先行研究[比企野 2011, 斎藤 2013]で開発された語りかけエージェントは、一問一答で質問を行い、高齢者に語りかけを行うものである。また頷きや相槌、「私は元気です」といったコメントを返す。図 1 にエージェントとの対話様子を示す。エージェントとの対話終了後、発話内容や発話時間などを記載した xml 形式のログファイルが生成される[酒井 2012]。



図 1 語りかけエージェントとの対話様子

3. アノテーションに基づく反応性の分析

3.1 反応性のアノテーション

高齢者の状態把握に向けて、まず反応性の違いが人の直感でどれだけ一致しているかを調査した。10名の認知症高齢者(男性2名、女性8名、平均年齢:77.6、平均MMSE:23.5)の対話の様子を記録したビデオデータを観察し、2名のアノテータによって各発話に対する反応性を4段階で分類してもらった。発話数は全部で162発話である。低い評価である1,2に分類された発話をL(Low)、高い評価である3,4に分類された発話をH(High)とラベルづけた。

2名のアノテータの分類結果を基に、偶然によらない一致率の指標である κ 係数を式(1)により求めた。

$$\kappa \text{ 係数} = \frac{\text{見かけ上の一致率} - \text{偶然による一致率}}{1 - \text{偶然による一致率}} \dots (1)$$

κ 係数の値は0.34687であり、2名のアノテータによる判断の一致率は低いという結果になった。

3.2 アノテーションデータの分析

κ 係数による一致率が高くなかったため、2名のアノテータの判定が一致しているデータのみを用いて、以下の5つの観点で分析を行った。

(i) 発話間のポーズ: エージェントの発話終了から被験者が発話を開始するまでの時間

- (ii) ピッチ: 被験者による発話区間のピッチ(声の高さ)の平均
 - (iii) 発話長: エージェントが発話をしてから、次にエージェントが発話をするまでの被験者の発話区間の総時間
 - (iv) 抑揚: 被験者の発話区間におけるピッチの最高値と最低値の差
 - (v) 傾き時間: エージェント発話中に被験者が傾いた時間
- 両者が H と判定した発話を HR, L と判定した発話を LR とする. 各情報の平均および t 検定の結果を表 1 に示す. ポーズ, ピッチ, 発話長, 抑揚に関して HR, LR 間で有意差が見られた. 一方, 傾き時間は有意差が見られなかった. この結果から, 被験者の発話韻律特徴に着目し発話の分類を行うことにした.

表 1 HR および LR の平均と有意性

特徴	HR の平均	LR の平均	t 検定の結果
ポーズ(秒)	0.98	1.57	t(95)=-2.82, p<0.05
ピッチ(Hz)	178.6	137.1	t(105)=2.99, p<0.01
発話長(秒)	4.77	1.60	t(105)=5.83, p<0.01
傾き時間(秒)	0.61	0.34	t(105)=1.86, p<0.1
抑揚	171.4	93.0	t(105)=3.90, p<0.01

4. クラスタリングによる反応性の分類

3 節で得られた傾向から, 有意差が見られたピッチ, ポーズ, 発話長, 抑揚の 4 つの発話韻律特徴を用いて発話の自動分類が可能か検討した. 本研究ではクラスタリング手法の 1 つである k-means 法を用いた.

4.1 認知症高齢者の発話分類

先行研究の語りかけエージェントを使用して, 新たに対話実験を行った. この実験によって, 15 名の認知症高齢者(男性 10 名, 女性 5 名, 平均年齢:75.3, 平均 MMSE:21.7)の音声データを収集した.

これらの音声データ(ログが生成されなかった 2 名を除外)と 3 章で用いたデータを用いてクラスタリングを行った. 各特徴量は尺度が異なるため, 発話ごとの各特徴量の値を, 特徴量ごとの分散値で割って正規化し, クラスタ数を 2 に設定して k-means 法を適用した. 分析は男女別に行った.

表 2 に認知症高齢者の女性における各クラスタの中心点を表す 4 つの特徴の値を示す. クラスタ b はクラスタ a と比べて, 発話長が長く, ピッチが高く, ポーズが短く, 抑揚が大きいという結果になった. 男性についても同様の傾向が得られることを確認した(表 3). すなわちクラスタ b に分類される発話は, ある程度の長さを持ち, 元気がよく, 質問にすぐに返答したりリズムカルに話をしている発話であり, このような発話は反応性の良い発話と解釈できる. これにより, 生成されたクラスタを用いて認知症高齢者の発話を反応性の良い発話とそうでない発話に分類できることが明らかになった.

4.2 認知症高齢者と健常者の発話分類

4.1 節の分析結果について, 認知症高齢者に加えて健常者のデータを合わせて同様の傾向が見られるかを調査した. ここでは, 3.1 節で反応性の違いを調べた 10 名の認知症高齢者のデータと, 新たに 9 名の健常者(女性 9 名, 平均年齢:77.7)のデータを収集したものをを使用した.

表 2 クラスタリング結果(認知症高齢者女性のみ)

クラスタ	発話長	ピッチ	ポーズ	抑揚
a	0.267	0.291	0.186	0.017
b	0.638	0.306	0.032	0.040

表 3 クラスタリング結果(認知症高齢者男性のみ)

クラスタ	発話長	ピッチ	ポーズ	抑揚
a	0.193	0.170	0.022	0.029
b	0.397	0.227	-0.103	0.045

表 4 に k-means 法を適応した各クラスタの結果を示す. 4 つの特徴について, 表 2 と同様の傾向(発話長:長, ピッチ:高, ポーズ:短, 抑揚:大)となることを確認した.

表 4 クラスタリング結果(健常者・認知症高齢者)

クラスタ	発話長	ピッチ	ポーズ	抑揚
a	0.226	0.197	1.305	0.012
b	0.476	0.220	0.882	0.032

4.3 クラスタリングに基づくユーザの分類

4.2 節の結果について, 各被験者の個々の発話がクラスタ a, クラスタ b どちらに分類されるか調査した.

図 2 に各クラスタに分類される発言の割合を認知症高齢者, 健常者に分けて示す. 認知症高齢者はクラスタ a に分類される発話の割合が多い. 一方, 健常者はクラスタ b に分類される割合が多いことが示された. これにより認知症高齢者は反応性の良くない発話の頻度が高くなり, また反応性の割合によって認知状態を判断できる可能性があることが示唆された.

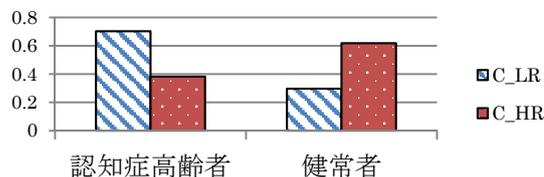


図 2 認知症高齢者/健常者の反応性割合

5. クラスタリングを用いた反応性の分類手法

表 2, 3 をもとに, 新たに入力された発話に対して反応性の良い発話かそうでない発話かの分類を行うことができる. 分類の手順を以下に示す.

① 新しく入ってきた発話の各特徴量(ピッチ, ポーズ, 発話長, 抑揚)を算出

② 算出された特徴量ごとに, 過去に行われた被験者の全発話の特徴量ごとの分散値で正規化

③ 正規化された特徴量(x_1, \dots, x_4)を 4 次元空間の点と考え, クラスタの中心点(c_1, \dots, c_4)との距離を式(2)より算出

$$\text{距離} = \sqrt{(c_1 - x_1)^2 + \dots + (c_4 - x_4)^2} \quad \dots (2)$$

距離の計算では, 表 2, 3 の性別ごとのクラスタを適用し, クラスタ a, b のうち距離がより近いクラスタに発話が分類される.

6. レポート用ダイジェスト文の生成

先行研究[比企野 2011, 斎藤 2013]で構築したエージェントとの対話後に出力される結果は xml 形式となっているため, 利用を想定している医師やご家族にとって非常にわかりにくいも

のとなっている。認知症高齢者との対話の様子を直感的に分かりやすい形で表現するために、対話を通した被験者の様子を簡潔にまとめたダイジェスト文を会話活動レポートの一部として生成する。

6.1 ポジティブな傾向を示す韻律情報の判定

ダイジェスト文を生成するために、表 1 で有意差の見られた 4 つの韻律特徴を利用する。ダイジェスト文生成の指針として、(I) 認知症高齢者全体の会話ログと比較し、他者と比べて良い反応が見られた特徴項目、および(II) 個人内の会話における各質問について、良い反応を示した特徴項目という 2 パターンについて、ポジティブな側面を要約した文章を生成する。

(I) の全体との比較では、ダイジェスト文生成対象となる会話でなされた全発話の特徴量の平均値を算出し、過去に実施された全被験者の平均値と比較して、その値が上回る特徴について言及する。(II) の個人内で比較をする場合は、各質問における発話の特徴量と、その会話全体の平均値を比較し、良い反応が見られた特徴を同定する。平均値との比較では、(I)、(II) ともに平均値以上、平均値+1σ 以下の場合には「良い(G1)」とし、1σ 以上の場合「とても良い(G2)」と判断する。

6.2 ダイジェスト文の生成手法

6.1 節で G1, G2 と判定された特徴に関して、あらかじめ定義されたテンプレートに基づきダイジェスト文を生成する。表 5 にテンプレートの一部を示す。テンプレートは全 23 種類定義されている。

まず、ダイジェスト文で言及される特徴の G1, G2 の個数に応じてテンプレートの候補を選択する。例えば G1 が 1 つの場合は、テンプレート T1 と T2 が候補となる。次に、各特徴が候補テンプレート中の変数部分の条件とマッチしているかをチェックする。例えば、T1 の変数部分の条件は、type1 発話長であるが、言及しようとする特徴が発話長であれば、条件にマッチする。その場合、表 6 中の発話長の type1 の値「たくさん話して」を変数部分に挿入し、ダイジェスト文を完成させる。

また、「G1: 発話長、抑揚、ポーズ」が言及される特徴である場合、T4 が適用され、「たくさん話してくれました。リズム、反応も良かったです。」というダイジェスト文が生成される。

表 5 テンプレートの例

	特徴の数	テンプレート
T1	G1:1 個	[type1 発話長] くれました。
T2	G1:1 個	[type1 ピッチ_ポーズ_抑揚] した。
T3	G2:1 個	とても[type1 発話長] くれました。
T4	G1:3 個	[type1 発話長] くれました。[type2 ピッチ_ポーズ_抑揚]、[type2 ピッチ_ポーズ_抑揚] も良かったです。
T5	G1:4 個	[type2all], [type2all], [type2all], [type2all] すべて良かったです。

表 6 ダイジェスト用の単語リスト

特徴	[type1]	[type2]
ピッチ	元気な声で	声の高さ
ポーズ	良い反応で	反応
発話長	たくさん話して	話した長さ
抑揚	良いリズムで	リズム

7. レポート生成システム

7.1 システムの構成

最終的な出力を統合し、レポートを表示するためのシステムを構築した。本システムの処理フローを図 3 に示す。

語りかけエージェントとの対話終了後に生成されるログ情報には、エージェントの発話終了時間や被験者の発話開始/終了時間、音声認識結果などがログファイル(xml形式)で保存されており、加えて対話時の各発話の音声ファイル(wav形式)、動画ファイル(flv形式)を保持している。システムでは、まず状態推定モジュールにおいて、音声ファイルを音声分析ソフトウェアの Praat を通して 4 種類の韻律特徴量を算出し、それに基づき各発話の反応性を 5 章の分類手法により「反応性の良い発話(good)」、「そうでない発話(not_good)」として判断する。また、ダイジェスト文生成モジュールでは、韻律特徴量の結果から、6.2 節で述べた手法に基づきダイジェスト文を生成する。レポート生成モジュールは、反応性の結果とダイジェスト文を受け取り、これらの情報を元のログファイルに書き加えた出力用のレポートファイルを生成する。

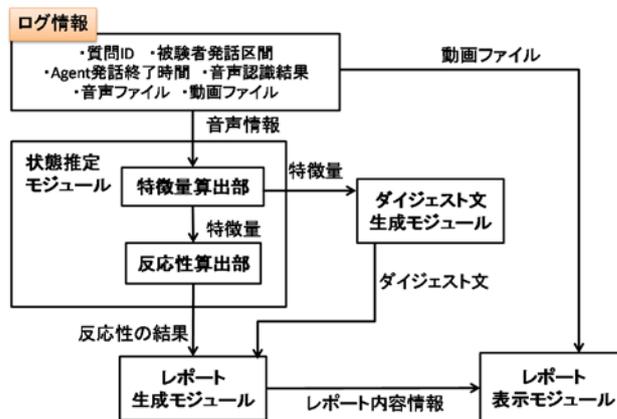


図 3 レポート生成システム

レポート表示モジュールでは、レポートファイルと対話時の動画ファイルを受け取り、認知症高齢者とエージェントとの対話の様子を GUI で表示する。システムはサーバ接続型の Adobe AIR アプリケーションとして MXML+ActionScript 言語で実装されており、ネットワークを利用できる環境であればどこからでも実行できる。サーバは、オープンソースで提供されたストリーミングサーバである Red5 を利用しており、保存された対話映像を接続先のクライアントに送信できるようになっている。

7.2 レポートの表示例

レポート表示画面を図 4 に示す。システムは以下の 4 つの項目から構成されている。

- ① ログファイル選択部
- ② ダイジェスト文表示部
- ③ エージェント質問・反応性表示部
- ④ ビデオ表示部

サーバが起動している状態で、①のエリアにサーバのホスト名とポート番号、表示したい会話ログファイルを入力し、接続ボタンを押すことでレポートが表示される。②のエリアでは、ダイジェスト文が表示される。ここでは、反応性の良い発話が全体の 6 割以上を占める場合、6.1 節で述べた (I) のダイジェスト文を、0.6 未満の場合、(II) の文をレポート結果として表示される。

③のエリアには、エージェントの質問に対する反応性の結果が、「good」「not_good」の判定結果に応じて笑顔/通常のフェ

イスマーク画像で表示される。各質問欄をクリックすることで、④のビデオ表示部にその質問終了時点の認知症高齢者の発話映像が再生される。

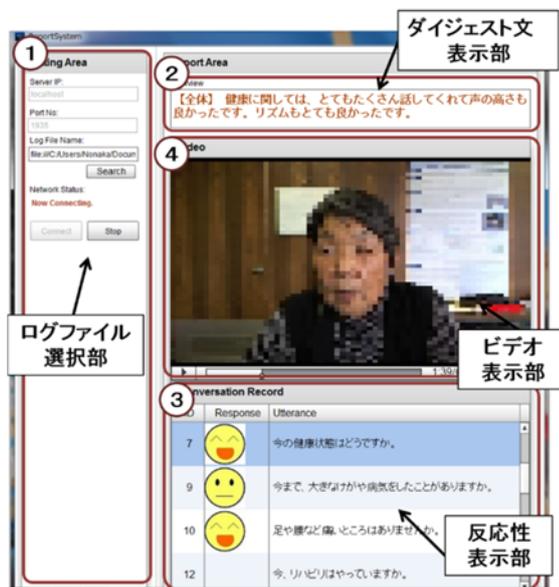


図4 レポート表示画面

8. 結論

本研究では、語りかけエージェントとの対話における反応性の違いを調査し、発話韻律情報による反応性の自動判定方式を提案した。また、発話韻律情報を用いたダイジェスト文の生成を行い、ダイジェスト文と反応性結果を使い医師やご家族に提示するための会話活動レポートの生成を行った。今後は、生成されたレポートやダイジェスト文の妥当性を医師や介護者に確認して貰うといった調査、また表情など他の情報を追加することでさらに性能の良いレポートを生成できるようにする。

参考文献

- [Pollack 2005] Pollack, M.: Intelligent technology for an aging population: The use of AI to assist elders with cognitive impairments, *AI Magazine*, Vol.26, No.2, pp.9-24, 2005.
- [酒井 2012] 酒井洋一, 野中裕子, 安田清, 中野有紀子: 認知症患者のコミュニケーション特性の評定における会話エージェントの利用: 第74回情報処理学会, 2012.
- [比企野 2011] 比企野純太, 中野有紀子, 安田清: 「会話エージェントを利用した認知症患者のためのコミュニケーション支援」, 第73回情報処理学会全国大会論文集, 4ZA-7, pp.195-196, 2011.
- [齋藤 2013] 齋藤直子, 林佑樹, 中野有紀子, 安田清: 認知症患者の状態把握支援のための発話内容認識手法, 第27回人工知能学会全国大会, 2G1-4, 2013.