

# 認知症患者の状態把握支援のための発話内容認識手法

## Automatic Analysis of Utterance Contents for Assessing the Status of Elderly People with Dementia

齋藤直子\*<sup>1</sup>  
Naoko Saito

林佑樹\*<sup>2</sup>  
Yuki Hayashi

\*<sup>1</sup> 成蹊大学理工学部/東京工業大学大学院総合理工学研究科  
Seikei University/Tokyo Institute of Technology

中野有紀子\*<sup>2</sup>  
Yukiko Nakano

安田清\*<sup>3</sup>  
Kiyoshi Yasuda

\*<sup>2</sup> 成蹊大学理工学部  
Seikei University

\*<sup>3</sup> 千葉労災病院/京都工芸繊維大学  
Chiba Rosai Hospital/Kyoto Institute of Technology

Recently, the number of people with dementia is increasing in our country. In order to support them, we have developed a conversational agent for elderly people with dementia. In this study, our objective is to analyze the cognitive status of the elderly people based on their utterances observed in conversations with the agent. Since care givers usually estimate mental and physical condition of elderly people based on some specific keywords in utterances, we focus on verbal information of elderly people. In order to detect topics of each utterance, we exploit Latent Semantic Analysis (LSA) technique which produces a set of concepts based on the input documents. In addition, for estimating the positive/negative emotion expressed in utterances, we propose an emotion estimation method based on Plutchik's emotion model and Japanese WordNet thesaurus. Experimental results showed that our method successfully detects the topics and emotions of each utterance.

### 1. はじめに

近年、日本では認知症患者の数が増加傾向にある。この傾向は社会の高齢化に伴い今後も続く予測される。しかし、認知症患者の介護や身の回りの世話をする介護・看護従事者の数はその需要の高まりに追いついておらず、負担が大きくなってきているのが現状である。

認知症には様々な症状があるが、実行機能障害や記憶障害などをシステムが認識するには、言語と非言語の両側面からのアプローチが必要である。現場における介護・看護従事者は、患者の状態を発話に含まれる心理や体調に関するキーワードから把握している。そこで本研究では、患者の発話情報に着目して、発話に含まれるキーワードからトピックと感情を抽出し、レポートに纏めることのできる発話内容認識手法を提案する。

### 2. 語りかけエージェントシステム

比企野ら[比企野 2011]は認知症患者との非課題遂行対話を行うエージェント(図 1)を構築した。対話を円滑に進めるには、エージェントのフィードバックが重要であるため、本エージェントは、適切なタイミングで相槌や頷きを行い、ユーザが話し終わってから次の質問に進むという協調的な対話機能を持つ。これらの機能を持たない非協調的エージェントとの比較実験では、対話の中で適切なタイミングで相槌を行うエージェントの有効性が示されている。酒井ら[Sakai 2012]は、同エージェントシステムに音声認識機能を追加し、ユーザ発話の音声認識結果をログファイルに残す機能を追加した。ここでは Google 音声認識を利用し、音声認識結果とその信頼度をログファイルに出力している。

本研究では語りかけエージェントと認知症患者の対話におけるログファイルを利用して、発話を分析する。



図 1 語りかけエージェントの外観

### 3. 健常者の発話データ収集実験

以降、認知症の診断を受けていない人を健常者と定義する。後述する LSA のコーパスに利用するため、健常者の発話データを収集する目的で実験を実施した。看護師が整形外科の入院患者に頻繁に行う質問を事前に調査し、それをもとにエージェントの質問内容を決定した。これは、看護師が認知症ではない高齢者に日常的に投げかける質問を認知症患者に投げかけたときに、健常者の応答と比較することで認知症患者に特有の兆候がみられるのではないかと考えたためである。

被験者には語りかけエージェントと 1 対 1 で 10 分程度の対話を行ってもらい、その対話を録音した。実験環境を図 2 に示す。被験者の前にはエージェントを出力するモニター、音声出力をするスピーカー、実験の様子を撮影するビデオカメラを設置し、エージェントと対話をして頂いた。対話の録音には、指向性の

高いピンマイクを使用し、録音の際の音量調節用にオーディオインターフェースを用いた。

本実験を通して、64～93歳(平均年齢77.7歳)の健常者9名(全て女性)のデータを収集した。

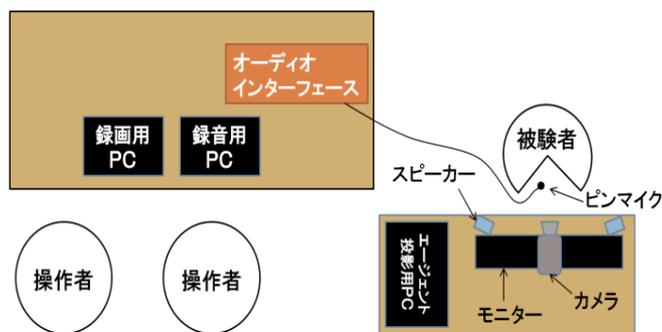


図2 健常者実験の環境

#### 4. 発話内容認識モデルの構築

認知症患者の特徴として、以下が挙げられる。

- (1) 認知機能低下による会話の不成立(相手と自分の発話トピックの不一致)
- (2) 不安や鬱などの精神症状

患者の発話からこれらの特徴を判定するため、患者の発話内容からトピックと感情を抽出する必要がある。(1)では患者の発話内容から抽出されたトピックとエージェントの質問トピックが一致しているかどうか重要であり、(2)では否定的(Negative)な感情を抽出することで不安や鬱状態を発見できる可能性がある。そこで本研究では、トピック推定と感情推定を行うモデルを構築する。

##### 4.1 トピック推定モデル

患者の発話のトピックを推定するために、本研究では潜在的意味解析(Latent Semantic Analysis: LSA) [Deerwester 1990]を用いる。潜在的意味とは、文書内の各単語の重みと言った単純な特徴ではなく、様々な統計的解析を行うことにより抽出される特徴である。LSAは、文書とその文書に含まれる単語を利用して単語文書行列を作成し、その行列に対して特異値分解と次元削減を行うことで潜在的意味の軸を抽出する手法である。次元削減により、高次元の空間では別々に扱われていた検索語が、低次元の空間では同じ潜在的意味を持つ可能性があるため、検索語の持つ意味や概念に基づく検索ができる。LSAよりも簡単な手法としてベクトル空間モデル[Salton 1983]があるが、ベクトル空間モデルでは次元数の増加に従い検索時間の増大などの問題や不要な検索語に因る精度低下といった問題がある。LSAではこのような問題を次元削減により解決できる利点があり、コーパスの量と質次第で精度を向上できる。また、LSAは他のトピック推定手法よりも、より人間に近い推定結果が出るという特徴がある。本研究では、医師や介護士等の人間から見て患者の発話がどのようなトピックであるかを判断する結果と同様な判定を自動的に抽出することを目指しているため、LSAの利用が適切であると考えた。

本研究ではLSAによって抽出された潜在的意味の各軸に名前を与え、ユーザの発話が各次元と類似するかを、コサイン類似度を計算することでトピック推定を行う。

エージェントの質問に対する健常者の応答のデータを質問に対する正常な発話とみなし、単語文書行列を作成した。ここでは、健常者実験で収集した発話の書き起こしデータを用いた。発話データでコーパスを作成するにあたり、トピック推定と感情推定に必要な情報として主語と述語が重要であると考えた。そこで、コーパスに入れる語を用言(動詞、形容詞、形容動詞)と体言(名詞)に限定した。ただし、名詞の中でも形態素解析エンジンのlucene-gosen<sup>1</sup>で代名詞・接尾・非自立語に分類されるものや、「ある」、「する」といったトピック推定と感情推定に重要でないと考えられる語も除外した。これは、単語文書行列の各要素には文書内の単語の出現頻度が入っているが、LSAでは文書間における単語の共起が影響することがコーパスを分析することによりわかったためである。出現頻度は高いが意味として重要でない語を除外することにより、エージェントの質問トピックにより近いトピックに分類できるようになると考えた。

単語文書行列の次元を増減させ、各次元に属する文書集合を検討した結果、手動でトピックに名前を与えるために最適な値として次元数を8個に設定した。設定したトピック名を表1に示す。これらのトピック名(7種類)をシステムのトピック推定部の出力とする。

表1 トピックの分類

次元 No.	トピック名
1	その他
2	住所
3	健康
4	生活
5	飲食
6	読書(視力)
7	健康
8	趣味・年齢

##### 4.2 感情推定モデル

不安や鬱という精神症状を発話から見つけ出すには、その発言が肯定的な感情(Positive)によるものなのか、否定的な感情(Negative)によるものなのかを発話ごとに推定できることが重要である。感情を推定するにあたり、PositiveまたはNegativeな感情を表す語を名詞、動詞、形容詞、形容動詞に限定して収集した。

Plutchikの感情の輪[Plutchik 1991]を参考に、PositiveワードリストとNegativeワードリストを作成した。Plutchik理論に基づく感情分類を、PositiveとNegativeに分類した結果を表2に示す。なお、本研究ではPlutchik理論においてsurprise軸と混合感情に分類されるものに関してはPositiveとNegativeという概念に合わない判断のため除外することとした。

<sup>1</sup> lucene-gosen: <http://code.google.com/p/lucene-gosen/>

表 2 Plutchik 理論による感情分類

感情の強度	低 ←————→ 高		
Positive Words	平穩	喜び	歡喜
	受け入れ	信頼	崇拜
	興味	期待	用心
Negative Words	心配	恐れ	恐怖
	落胆	悲しみ	悲嘆
	退屈	嫌悪	憎悪
	苛立ち	怒り	激怒

また、Plutchik の感情モデルに存在しない日本語の感情語があることを考慮して追加ワードを独自に設定し、Positive / Negative ワードリストに追加した。ここでは、各語に感情の強度を踏まえた重みを付与する。表 2 の中央のカラムに属する感情は基本感情と呼ばれる。基本感情より感情の強度が高いものには  $+\alpha$ 、低いものには  $-\alpha$  の重みを付与した。なお、 $\alpha$  は経験的に定めた値である。各語に付与した重みを表 3 に示す。

表 3 各語の重み ( $\alpha=0.1$ )

語の種類	重み
Plutchik の基本感情 (表 4 の中央カラムにある語)	1.0
基本感情より感情強度が高い語	基本感情の重み(=1.0) $+\alpha$
基本感情より感情強度が低い語	基本感情の重み(=1.0) $-\alpha$
追加ワード	1.0

次に、リスト内の各語について日本語 WordNet<sup>2</sup>を利用して同義語を抽出した語を更にリストに追加した。抽出された同義語の意味は、元の語の意味とわずかに差が生じる恐れがある。そこで、WordNet 利用による元の語との意味の距離を考慮し、ここでは(元の単語の重み -  $\beta$ )の重み付けを行った( $\beta=0.2$ )。なお、 $\beta$  は経験的に定めた値である。

この重みを利用して、発話の感情推定を行う。ある発話  $u$  に含まれる Positive word リストの語の集合を  $W_{u,p}$ 、Negative word リストの語の集合を  $W_{u,n}$  とし、各リストの語  $i$  の重みを  $w_i$  としたとき、発話  $u$  の感情得点  $P_u$  を式(1)により算出する。

$$P_u = \sum_{i \in W_{u,p}} w_i - \sum_{j \in W_{u,n}} w_j \quad (1)$$

上式で計算した感情得点に基づき、発話  $u$  の感情  $F_u$  を Positive, Neutral, Negative のいずれかとして式(2)より判定する。

$$F_u = \begin{cases} \text{Positive} & (P_u > 0) \\ \text{Neutral} & (P_u = 0) \\ \text{Negative} & (P_u < 0) \end{cases} \quad (2)$$

この Positive, Negative, Neutral と感情得点を感情推定モジュールの出力結果とする。

## 5. 発話内容認識システムの実装

4 章で用いたモデルを用いて、発話内容認識システムを実装した。システムの構成を図 3 に示す。ユーザ発話の音声認識結果を[Sakai 2012]のシステムのログファイルとして受け取る。ここでは音声認識結果のうち信頼度第 1 位の結果をユーザの発話とし、形態素解析 lucene-gosen にかけて形態素単位に分割した結果に基づき、感情推定モジュール及び、トピック推定モジュールで処理を行う。本システムは、各モジュールから出力された結果をレポート形式でまとめたファイルを出力する。レポートの出力例を表 4 に示す。

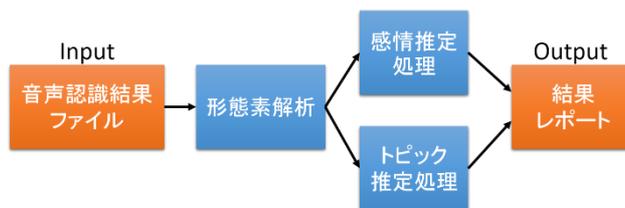


図 3 システム構成図

表 4 レポートの出力例

エージェント質問	ユーザ応答の書き起こし	感情	感情得点	トピック
ご飯は食べてますか	はい、食べてます	Neutral	0	飲食
趣味はなにかありますか	お花を作ることが趣味です	Positive	0.7	趣味・年齢
今まで、大きなけがや病気をしたことがありますか	特にありませんが、物忘れがあるようなので心配です	Negative	-0.9	生活

システムは語りかけエージェントの音声認識結果を適用できる仕様となっているが、表 4 では実際の認知症患者の発話の書き起こしを入力とした出力例を示している。このように 1 組のエージェントの質問と患者の応答に対する、感情、感情得点、及びトピックを出力することで、患者の応答の特徴を表示できる。

## 6. 評価実験

### 6.1 実験設定

5 章で実装した発話内容認識システムについて、トピック推定モジュールと感情推定モジュールの精度を検証するため、被験者 5 名に認知症患者の発話 50 個、健常者の発話 50 個の計 100 個の発話に対するトピックと感情を選択してもらった。ここでは、各発話に対して、感情を -3 (Negative) ~ 0 ~ +3 (Positive) の 7 段階の数値を付けてもらい、トピックを表 1 の 7 種類から選別してもらった。なお、本システムは音声認識結果を入力として発話を分析する構成となっているが、本実験では音声認識の精度に影響されることなく各モジュールの精度を評価するために、発話データの書き起こし文をシステムの入力とした。

<sup>2</sup> 日本語 WordNet: <http://nlpwww.nict.go.jp/wn-ja/>

感情推定の評価では、被験者に付けてもらった感情の数値の平均を  $p$ 、システムが出力した感情得点を  $s$  としたとき、表 5 に示す尺度に基づき  $p$  と  $s$  の一致率を精度とする。また、トピック推定の評価では、システムが検出したトピックが、5 名の被験者のうち何名の被験者が付けたトピックと一致したかを発話ごとに算出し、その平均値を精度とする。

表 5 感情推定の評価尺度

感情分類	Negative	Neutral	Positive
$p$ の値	$p \leq -1$	$-1 < p < 1$	$1 \leq p$
$s$ の値	$s < 0$	$s = 0$	$0 < s$

## 6.2 実験結果

実験結果を表 6 に示す。表に示されるように、患者、健常者の両発話について、システムは約 7 割の精度で被験者の付けた感情を正しく検出できた。また、被験者が選択したトピックとシステムが推定したトピックの一致率は約 5 割であり、ある程度正しく発話のトピックを検出することに成功した。全体的にみると、今回は認知症患者の発話に対してより高い精度で推定ができていた。これは、認知症患者の方が発話の語彙が少なく、全体に短い発話が多かったことが影響していると考えられる。

表 6 感情推定とトピック推定の結果

評価項目	患者	健常者	全体
感情推定の精度	0.76	0.66	0.71
トピック推定の精度	0.57	0.45	0.51

実験結果では、「悲しくない」という発話を被験者が Positive と判断している一方で、システムは Negative と判断していた発話が見られた。本研究では、感情推定モデルを構築する際に「悲しくない」、「悲しくありません」といった否定形を考慮していない。係り受け解析を用いて肯定／否定表現を判別できるモデルに拡張することで、このような誤認識を改善できる見込みがある。また、本研究では名詞・動詞・形容詞・形容動詞を対象として感情推定をしているが、笑いを表す語（「ははは」、「うっふっふ」など）について被験者は高い数値を付ける傾向が見られた。笑い声に関する言葉は人に因って様々であるため、言語情報から笑い声を検出することは困難であるが、音声データを分析する際に得られるピッチ変化の特徴から笑い声を検出できると考えられる。このように言語情報に加えて非言語情報を考慮したモデルに拡張することで、今後感情推定の精度を向上できる可能性がある。

トピック推定に関して患者データと健常者データを比較した結果、健常者の方が被験者の付けたトピックが分散するという傾向が見られた。これは健常者の方が長い発話が多く、トピックが 1 つに絞れない発話が多かったためと推測できる。また本研究では、トピック推定部で発話と最も類似度の高い次元に付けられたトピック名を出力しているが、全ての次元の類似度が低く算出される発話も存在していた。現時点ではその中でもっとも類似度が高い次元のトピック名を出力しているため、全ての類似度が低い発話に関してはトピックを「その他」として判定することで、より正確にトピックを推定できる可能性がある。

提案したモデルでは、発話の特徴づける語が LSA のコーパスや各ワードリストに含まれるかどうかで結果が左右されてしまうため、今後も継続して多くの発話データを集め、LSA のコーパス拡張や語の重み付けの方法を検討していく必要がある。

## 7. まとめ

本研究は認知症患者の発話データを入力として感情推定とトピック推定を行い、結果をレポートとして出力する発話内容認識手法を提案した。提案手法を基に実装したシステムの評価実験を行い、発話の書き起こしデータを用いた場合、感情推定モジュールでは約 7 割、トピック推定モジュールでは約 5 割の精度で、被験者が発話を分類した感情、トピックを推定することができた。

現在、音声認識結果を入力としたシステムの出力を評価するために認知症患者の対話データ収集実験を行なっている。今後、このデータを用いてモデルを評価し、より正しく感情・トピックを推定できるようにコーパスを拡張していく予定である。また、本研究が対象としている言語情報に加えて傾きや音声のピッチといった非言語情報も扱うことで、様々な観点から状態把握を行うことができるシステムを目指していく。

## 参考文献

- [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, 4.195-4.196, 2011.
- [Sakai 2012] Yoichi Sakai, Yuuko Nonaka, Kiyoshi Yasuda, Yukiko I. Nakano: “Listener agent for elderly people with dementia”, Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction (HRI2012), pp.199-200, 2012.
- [Deerwester 1990] S. Deerwester: “Indexing by Latent Semantic Analysis”, Journal of the American Society for information Science, Vol.41, Issue 6, pp.391-407, 1990.
- [Plutchik 1991] R. Plutchik: “The emotions”, University Press of Amer, 1991.
- [Salton 1983] G. Salton and M. J. McGill: “Introduction to Modern Information Retrieval”, McGraw-Hill, 1983.