

雑談対話の語彙と係り受け情報を用いた認知症傾向検出

情報科学科 石原 颯人

指導教員：入部 百合絵

1 はじめに

2012年時点における65歳以上の高齢者の4分の1が認知症またはその予備軍であるとされ、今後も認知症高齢者は増加し続けることが予想される[1]。一方で、認知症は早期に治療を始めることで、症状の進行を遅延、抑制することができると言われており、認知症の早期発見は課題である。しかし、65歳以上の高齢者の夫婦のみの世帯、単独世帯の数は増加傾向にある。そのため、認知症傾向の発見が遅れ、認知症の症状が悪化してしまうケースも少なくない。そのような背景の下、病院に行かなくとも、日常生活の中で容易に認知症傾向を早期発見できることが期待されている。

田中ら[2]や柴田ら[3]は、認知機能テスト時の音声や映像、画像の内容を説明させる音声を利用した認知症の検出を行っている。しかし、テスト時発話は普通の会話とは話し方や内容が異なる。また、テスト自体に不快感を示す被験者も少なくない。

本研究では、普段の生活の中で比較的簡単に情報収集が可能であり、かつ認知症の症状が現れやすいと考えられる雑談対話中の音声に着目する。本稿では対話音声に含まれる言語情報から認知症の症状を捉えた語彙と係り受け情報を抽出することで認知症傾向の検出を目指す。

2 収集した雑談対話データ

本研究で使用したデータは徳島にある高齢者施設において、70歳以上の高齢者24名との雑談対話音声である。認知症傾向の判断は収録時に行ったHDS-R(改訂長谷川式知能評価スケール)の結果を利用し、30点満点中20点以下を認知症の疑いありと判断する。被験者の認知症傾向は傾向ありが13名、傾向なしが11名であった。

3 抽出した語彙と係り受け情報

24名の対話音声の書き起こし文を形態素解析MeCabで解析し、語彙と係り受け構造を中心に認知症の症状を捉えるための特徴量を抽出し、語彙に関する特徴量としてTTR(Type Token Ratio)、異なり名詞割合、日本語教育語彙レベル、代名詞割合、一般名詞割合と各品詞の割合を抽出した。日本語学習語彙レベルは、日本語教育学習語彙表で定義される難度に基づき抽出を行った。

前頭側頭葉変性症などの認知症においては文法誤りや複雑な構文の利用頻度が減るなどの症状が現れるとされる。本研究では特に複雑な構文に着目し、その特徴を捉えるために係り受け情報を抽出する。係り受けとは文節と文節の関係を示したもので、一般的に複雑な文章で長い係り受けが発生するとされる。係り受け距離の算出には係り受け解析器CaboChaを利用した。まず、各発話における係り受けの最大距離をその発話の最大係り受け距離として、被験者ごとに全ての発話について最大係り受け距離を求めた。さらに、各被験者が平均的にどの程度複雑な文章を利用しているかを求めるために、被験者ごとに全発話の最大係り受け距離の平均を求め、各被験者の平均最大係り受け距離として抽出した。ただし、「はい」という返事のみなど、係り受けが発生しない発話は分析対象から除いた。

Table 1 認知症傾向有無に対する平均最大係り受け距離の比較

	平均値		t検定 p値
	傾向あり	傾向なし	
雑談対話	2.6	4.5	0.012*
HDS-R時	1.6	2.5	0.029*

p<0.05:*

Table 2 雑談対話とHDS-R実施時の音声を用いた識別結果

係り受け	雑談対話		HDS-Rの回答音声	
	利用なし	利用あり	利用なし	利用あり
正解率	0.75	0.96	0.88	0.88

抽出した平均最大係り受け距離について、t検定を行った結果をTable 1に示す。傾向有無間で5%水準の有意差が確認され、認知症傾向と平均最大係り受け距離に関係があることが示唆された。

4 語彙と係り受けを利用した識別結果

10分割交差検証でSVM(Support Vector Machine)により識別を行った結果をTable 2に示す。複数の識別器を比較した結果、SVMの結果が最も高かった。最大係り受け距離を利用しない場合、HDS-R実施時の音声を利用した方が、雑談対話の識別結果よりも正解率が高かった。一方で、特徴量に係り受け距離を追加した結果については、雑談対話音声を利用したときの識別精度が90%を超え、HDS-Rの音声を利用した場合の結果よりも高い結果となった。雑談対話は自由発話であるため、発話文の複雑さに差が生じやすくなり、雑談対話の方が優位な結果を得られたのではないかと考える。

5 おわりに

本研究では語彙と係り受け構造に着目し、高齢者の雑談対話から認知症傾向の検出を試みた。複雑な構文を示すとされる係り受け情報を用いた結果、認知症傾向の有無間で平均最大係り受け距離に有意差が認められたため、認知症傾向の検出に係り受け距離を利用した。その結果、検出の正解率が20%以上向上し、96%という高い結果を得ることができた。このことから、雑談対話の係り受け情報は認知症傾向の検出に有用であることが示された。

今回は被験者数も対話数も少数であったため、引き続き雑談対話の音声を収録するとともに、雑談対話音声からの認知症傾向検出を行っていく。

参考文献

- [1] 内閣府, "令和元年版高齢社会白書"(2019)
- [2] 田中宏季他, "アバターとの対話によるマルチモーダル情報を伴った早期認知症の検出", 情報処理学会研究報告, 14, pp.1-4 (2016)
- [3] 柴田大作他, "対照群付き高齢者コーパスの構築とそれを用いた認知症予備軍スクリーニング技術の開発", 人工知能学会論文誌 34 巻 4 号 B1-9(2019)