

近時記憶課題と人型ロボットとの日常会話における 軽度認知障害患者の発話特徴分析

吉井謙太¹ 木村大毅² 小杉晋央² 新川香² 高瀬 俊郎² 小林 正朋²
山田康智² 根本 みゆき³ 渡辺 亮平³ 塚田 恵鯉子³ 太田 深秀³
東 晋二⁴ 根本 清貴³ 新井 哲明³ 西村雅史¹

概要：認知症の早期発見を目指し、認知機能検査時の音声に含まれる特徴から認知機能低下の兆しを見出そうとする研究がこれまで数多く行われてきた。認知症の前駆段階である軽度認知障害(MCI)は近時記憶能力低下と関係があることが知られている。近時記憶に関する話題に着目することで MCI 患者を発見できる可能性がある。一方、介護現場ではロボットを用いた見守りやロボットが人の代わりに高齢者の話し相手となることが期待されており、ロボットとの日常会話から認知機能低下の兆しを見出すことができれば、特別な検査を行うことなく認知症早期発見に繋げられる可能性がある。本研究では、健常者 46 名と MCI 患者 46 名を対象に発話特徴分析を行った。認知機能検査において近時記憶能力を測る遅延再生課題音声を収録し、9 種類の特徴量を抽出し健常者と MCI 患者間の発話特徴に違いがあるか分析を行った。その後、同じ実験参加者から収録した人型ロボットとの日常会話音声のうち、近時記憶に関する質問を含んだ食事の話題の音声から特徴量を抽出し、同様に発話特徴分析を行った。結果、遅延再生課題音声では 6 種類、人型ロボットとの日常会話音声では 5 種類の特徴量で有意差と有意傾向が得られた。近時記憶に着目した課題や話題の音声から抽出した特徴量を用いることで、認知機能検査得点では違いが無い実験参加者や日常会話音声から MCI 患者を早期発見できる可能性が得られた。

1. はじめに

超高齢社会の日本において、認知症は深刻な社会問題である。認知症は早期発見し治療を開始することにより回復や症状の進行抑制が可能である。このため、認知症は早期発見が重要である。特に、認知症の前駆段階である軽度認知障害(Mild Cognitive Impairment)[1]の発見が重要であると言われている。認知症は MMSE (Mini Mental State Examination) [2]といった認知機能検査や MRI、血液検査などから総合的に判断し診断が下される。これらの検査は医療機関への受診が必要であり、時間や費用、人手と行った制約から気軽に行えるものではない。認知症早期発見のために、日常的に利用可能な認知症簡易スクリーニングが求められている。

認知症簡易スクリーニングとして、鳥羽[3]は、認知機能検査のうち、即時記憶を測る復唱と近時記憶を測る遅延再生によって、簡単におおよその知能を測定する方法を提唱した。認知症の前駆段階である軽度認知障害においても近時記憶能力の低下が見られることがあると知られているが、鳥羽は健常者と同様に軽度認知障害患者(以下、MCI 患者)も遅延再生課題達成が可能であると述べており、遅延再生課題の可否だけの判別は難しい可能性がある。

より簡易な方法として検査ではなく、会話音声から認知症や軽度認知障害の兆候を見出そうとする研究が数多く報告されている[4][5][6][7]。しかし、これらの多くは認知機能検査中の会話音声を対象としたものであるため、日常的な利用には適さない。一方、人手不足が深刻な介護現場では

ロボットを用いた見守りや高齢者の話し相手になることでの省力化が期待されている[8][9]。ロボットと高齢者の日常会話から認知機能低下の兆しを見出すことができれば、特別な検査を行うことなく会話を楽しんでもらいながら、日常的に認知症簡易スクリーニングを行える可能性がある。

先に我々は人型ロボットとの日常会話を用いた認知症簡易スクリーニングの基礎的検討として、認知機能検査である MMSE の見当識課題実施時の収録音声と人型ロボットとの日常会話音声から、それぞれ特徴量を抽出し発話特徴分析を行った結果、健常者と MCI 患者間で発話特徴が異なると示した。さらに、見当識課題音声と人型ロボットとの日常会話音声の比較分析から、人型ロボットとの日常会話音声を用いた認知症簡易スクリーニングの可能性を示した[10]。しかし、人型ロボットのどのような話題が MCI 患者を発見するのに有効か分かっていない。そこで本研究では鳥羽の知見を参考に実験参加者(健常者、MCI 患者)と人型ロボットの会話音声のうち、近時記憶に関する質問を含む食事の話題に着目した。食事の話題から特徴量を抽出し、健常者と MCI 患者間の発話特徴の相違点を分析し、近時記憶に着目した話題から MCI 患者発見の可能性を検討する。なお、近時記憶が MCI 患者発見に有効であるか検討するために、専門の心理士が実施した認知機能検査(MMSE)のうち即時記憶を評価する 3 単語の復唱課題と近時記憶を評価する 3 単語の遅延再生課題の 2 種類の音声を対象とし、健常者と MCI 患者の発話特徴の差異を分析した。

1 静岡大学大学院 総合科学技術研究科
Shizuoka University
2 IBM 東京基礎研究所
IBM Research Tokyo

3 筑波大学
University of Tsukuba
4 東京医科大学茨城医療センター
Tokyo Medical University Ibaraki Medical Center

2. 会話データ収録

2.1 認知機能検査

国際的に広く利用されている認知機能検査である MMSE を専門の心理士が実施した際の会話データを収録した。MMSE は見当識、記憶力を口頭で問う質問や、図形模写といった筆記を含む 11 問で構成された検査である。

今回は、鳥羽[3]の手法を参考に MMSE における 3 単語の復唱課題と 3 単語の遅延再生課題を分析対象とした。3 単語の復唱課題は、心理士が口頭で「ボール・はた・さくら」など 3 つの単語を教示し、実験参加者に復唱をしてもらって即時記憶を測る課題である。計算課題などを挟んだ後、3 単語の遅延再生課題として心理士が「先ほどの覚えてもらった 3 つの単語は何でしたか、教えてください」と質問し、復唱課題で教示した単語を思い出してもらうことで、近時記憶を測る。どちらも口頭で回答し、1 つの単語が正解するたびに 1 点が与えられる 3 点満点の課題である。



図 1 人型ロボットとの会話音声収録の様子 図 2 咽喉マイクと集音マイク

2.2 人型ロボットとの日常会話

ソフトバンクロボティクス社の人型ロボット「Pepper」[11]を用い、人型ロボットとの日常会話音声の収録を行った。なお、収録方法は文献[10]と同様であるが健常者、MCI 患者各 1 名ずつ実験参加者が増加している。ロボットとの会話データ収録の様子を図 1 に示す。実験は、Wizard of Oz の形式で実施した。オペレーターは実験参加者から見えない位置に座り、事前に用意したシナリオを基本としつつ、実験参加者の応答に応じて話題の遷移や発話内容の選択を適宜行った。

シナリオは質問応答パートと雑談パートで構成されており、質問応答パートではロボットが実験参加者に対して質問を投げかけ、それに対して実験参加者が答える 1 問 1

答形式とした。質問応答パートには、昨晚の食事内容や作り方など、記憶や論理的思考力を確かめる質問を織り交ぜた。一方、雑談パートでは、日常会話を模擬するために実験参加者の応答に適した返事や自己開示を行った。なお、実験参加者がロボットに対して質問をしてきた場合は、「その質問にはお答えできません」といった返事や次の話題に移り、はぐらかすこととした。今回は、「昨日は何を食べましたか？」という近時記憶に関する質問が含まれる食べ物の話題を分析対象とした。

2.3 収録機材

認知機能検査の際には心理士、実験参加者両者に、人型ロボットとの日常会話では実験参加者にネックバンド型の咽喉マイクと集音マイクを首に同時装着(図 2)してもらい会話データの収録を行った (Linear PCM 形式/44.1k/ステレオ)。集音マイクは対象話者の口元から遠く離れた位置に居る人の発話や環境音も収録してしまう。このため発話重量や大きな外部環境音発生した場合に正確な対象話者の発話区間推定が困難となる。そこで、本研究では上記の問題点を考慮し、外部環境音に頑健な咽喉マイクを併せて使用した[12]。咽喉マイクを用いることで、対象話者の発話タイミングおよび発話区間などの正確な分析が可能となる。

2.4 実験参加者

実験参加者のうち認知機能検査や血液検査など様々な検査の結果から医師が健常と診断した 46 名、軽度認知障害 (MCI) と診断した 46 名、合計 92 名を対象に認知機能検査音声と人型ロボットとの日常会話音声の分析を行った。本研究では認知症早期発見を目標としているため、認知症の前駆段階である MCI 患者に絞り分析を行った。分析対象とした実験参加者の属性を表 1 に示す。

3. 発話特徴分析

3.1 分析特徴量

健常者と MCI 患者の発話特徴の差異を分析するため、実験参加者の会話データから特徴量を抽出し、発話特徴分析を行った。発話特徴分析に使用した特徴量を表 2 に示す。

まず咽喉マイクで収録された会話データを用いて発話区間とフィラーを人手でラベル付した。このラベルを用いて発話時間長からフィラー割合までの 7 種類の時間に関する特徴量を抽出した。咽喉マイクを用いることで、発話重量が発生した際に集音マイクでは正確な抽出が難しい反応時

表 1 実験参加者

	性別			年齢			3 単語の復唱課題得点 [点]			3 単語の遅延再生課題得点 [点]		
	男性	女性	合計	平均	標準偏差	範囲	平均	標準偏差	範囲	平均	標準偏差	範囲
健常者	20	26	46	70.6	4.9	61-80	3.0	0.2	2-3	2.7	0.6	1-3
MCI 患者	26	20	46	74.3	5.1	64-87	2.9	0.4	1-3	2.0	1.0	0-3

間といった特徴量の正しい抽出が可能となった。

次に、集音マイクで収録された会話データから2種類の音響特徴量を抽出した。咽喉マイクは環境雑音に対して頑健であるが、皮膚伝道音であるため通常の音響マイクとは特性が大きく異なる。このため、音響特徴量は咽喉マイクで得られた発話区間を参照し、集音マイク側データの当該部分から Praat[13]を用いて推定することとした。

表2 特徴量

特徴量名	概要
回答時間長 [秒]	1 ターン中の実験参加者の発話開始点から終了点までの時間長
反応時間長 [秒]	心理士または人型ロボットの発話終了点から実験参加者の発話開始点までの時間長
発話時間長 [秒]	回答時間長における発話時間の合算値
無音時間長 [秒]	回答時間長における無音時間(300ms以上)の合算値
無音時間割合[割合]	発話時間長における無音時間長が占める割合
フィラー長 [秒]	発話時間長における「えー」など、間を埋める発話時間の合算値
フィラー割合[割合]	発話時間長におけるフィラー長が占める割合
Jitter (local)	声の高さの揺れ
Shimmer (local)	声の大きさの揺れ

3.2 統計的有意差検定

表2に示した特徴量において、健常者とMCI患者間の統計的有意差検定を実施した。検定には Welch の t 検定を用いた。本研究では、p 値が 0.05 未満の場合に有意差ありとする。また、同時に識別力を評価するために効果量[14]を用いる。今回は、t 検定の代表的な効果量である Cohen's d を用いた。効果量の目安としては、0.20(Small), 0.50(Medium), 0.80(Large)を用いた。

表3 復唱課題音声の発話特徴分析結果

特徴量名	健常者	MCI 患者	P value	Cohen's d
	Mean (SD)	Mean (SD)		
回答時間長	2.35 (1.25)	2.64 (2.05)	0.41	0.17
反応時間長	0.58 (1.08)	0.84 (0.67)	0.17	0.28
発話時間長	1.97 (0.57)	1.96 (1.05)	0.96	0.01
無音時間長	0.39 (0.85)	0.68 (1.18)	0.17	0.28
無音時間割合	0.11 (0.15)	0.19 (0.22)	0.04 *	0.43
フィラー長	-	-	-	-
フィラー割合	-	-	-	-
Jitter (local)	0.03 (0.01)	0.03 (0.01)	0.96	0.01
Shimmer(local)	0.10 (0.02)	0.12 (0.03)	0.06 ·	0.39

· p < .1 * p < .05 ** p < .01 *** p < .001

4. 認知機能検査における発話特徴分析

4.1 復唱課題音声の発話特徴分析

4.1.1 結果

復唱課題音声においては、無音時間割合で有意差が得られ、Shimmer(local)で有意傾向を得られた。どちらも効果量の基準の 0.2(Small)を上回った。反応時間長、無音時間長においては、有意差が得られなかったものの効果量の基準は上回った。なお、フィラーは、今回分析した実験参加者においては観測されなかった。

4.1.2 考察

復唱課題は心理士が3単語を教示した後、すぐに単語を繰り返すという課題である。健常者とMCI患者間の得点で t 検定を行ったところ、 $t(72) = 0.71, p = 0.47$ で有意差は得られず、得点において2群間に差があると言えない。無音時間割合において有意差を、Shimmer(local)において有意傾向が得られたことにより、得点では見出せない認知機能低下の兆しを特徴量から見出せる可能性が得られた。しかし、有意差を得られたのは調査した9種類中1種類の特徴量のみであり、即時記憶能力を測る復唱課題から認知機能低下の兆しを見出すのは難しいと言える。

4.2 遅延再生課題音声の発話特徴分析

4.2.1 結果

遅延再生課題音声においては、回答時間長、無音時間長、無音時間割合、フィラー長、フィラー割合、Shimmer(local)において有意差が得られ、Jitter(local)では有意傾向が得られた。効果量では、無音時間割合で基準の 0.8(Large)を、Shimmer(local)では 0.5(Medium)を上回り、その他の有意差、有意傾向が得られた特徴量においては 0.2(Small)の基準を上回った。

4.2.2 考察

遅延再生課題は、復唱課題で教示された3単語を思い出し回答する課題である。遅延再生課題の得点で t 検定を行ったところ、 $t(73) = 3.98, p < .001$ となり、得点においても健常者、MCI患者間に有意差が得られた。MCI患者は遅延再生課題において失点しており、近時記憶が低下傾向にあ

表4 遅延再生課題音声の発話特徴分析結果

特徴量名	健常者	MCI 患者	P value	Cohen's d
	Mean (SD)	Mean (SD)		
回答時間長	6.66(10.71)	12.22(14.61)	0.04 *	0.43
反応時間長	0.88 (2.43)	0.45 (2.93)	0.24	0.25
発話時間長	2.91 (2.87)	3.71 (3.52)	0.24	0.25
無音時間長	3.60 (7.92)	7.94(11.55)	0.04 *	0.44
無音時間割合	0.26 (0.28)	0.47 (0.26)	<.001 ***	0.80
フィラー長	0.15 (0.69)	0.54 (1.01)	0.03 *	0.45
フィラー割合	0.01 (0.06)	0.02 (0.05)	0.34	0.20
Jitter (local)	0.03 (0.01)	0.04 (0.01)	0.08 ·	0.36
Shimmer(local)	0.12 (0.03)	0.14 (0.04)	0.02 *	0.52

· p < .1 * p < .05 ** p < .01 *** p < .001

ると示した。MCI 患者の無音時間長やフィルター長が健常者に比べ増加したのは、近時記憶の低下により単語を思い出すのに時間を要したためであろう。また、発話時間長において有意差が得られなかったにも関わらず回答時間長において有意差が得られたのは、無音時間長やフィルター長の増加が影響したと考えられる。Jitter(local)では有意傾向、Shimmer(local)においては有意差が得られたのは、MCI 患者が思い出しながら回答したことで声の高さや大きさの揺らぎが健常者より多く出たのであろう。

4.3 遅延再生課題における満点取得者の発話特徴分析

4.3.1 復唱及び遅延再生課題の得点分布

遅延再生課題の得点において、健常者に比べ MCI 患者の平均得点が低く、近時記憶が低下していると示した。t 検定は平均値の比較であるため、どの程度の実験参加者が失点しているか確認を行った。得点と人数の分布を表 5 に示す。

表 5 復唱課題と遅延再生課題の得点分布

得点	健常者[人]		MCI 患者[人]	
	復唱	遅延再生	復唱	遅延再生
3 点	44 ▶	36	43 ▶	20
2 点	2 ▶	7	2 ▶	11
1 点	0 ▶	3	1 ▶	12
0 点	0 ▶	0	0 ▶	3

健常者においては、46 名中 10 名と約 22%が遅延再生課題で失点しているのに対し、MCI 患者は 46 名中 26 名の約 57%が失点している。仮に遅延再生課題において失点した実験参加者を MCI 患者と分類する場合、2 割の健常者を誤って MCI 患者としてしまうのに対して、遅延再生課題で満点を取得している約 4 割の MCI 患者を誤って健常者と分類してしまう。MCI 患者を誤って健常者と判断してしまうのは、認知症の前駆段階である軽度認知障害を見落としていることを意味し、問題である。遅延再生課題で満点を取得しており、点数では違いがわからない実験参加者においても、特徴量で違いが出る可能性がある。そこで、遅延再生課題で満点を取得した実験参加者のみに絞り、表 2 で示した特徴量において発話特徴分析を実施した。なお、健常者と MCI 患者の満点を取得している人数に差があるため、今回は人数が少ない MCI 患者 20 名に健常者の数も揃え、発話特徴分析を実施した。また、4.1.2 において、復唱課題から認知機能低下の兆しを見出すのは難しいと考察したため、ここでは遅延再生課題のみを対象とした。

4.3.2 結果

遅延再生課題において満点を取得した健常者 20 名、MCI 患者 20 名の分析結果を表 6 に示す。無音時間割合のみで有意差が得られ、効果量においては 0.5(Medium)の基準を上回った。

4.3.3 考察

遅延再生課題の満点取得者では、無音時間割合のみで有意差が得られた。満点を取得しており、認知機能に問題ないように見える実験参加者においても無音時間割合から認知機能低下の兆しを見出せる可能性が示された。無音時間割合は表 4 に示した遅延再生課題の分析結果においても有意差が得られており、MCI 患者を見出すのに有効な特徴量の可能性がある。さらに、調査した 9 種類の特徴量のうち 6 種類の音響特徴量の効果量で目安となる基準を上回った。今回は健常者、MCI 患者各 20 名と分析対象が少なく、更なる分析が必要であるものの、遅延再生課題の得点において差が無い実験参加者に対しても、特徴量を用いることで MCI 患者を見出せると示唆された。

5. 人型ロボットとの日常会話音声における発話特徴分析

5.1 人型ロボットとの日常会話音声の分析対象話題

ここまでは、近時記憶に関する課題である遅延再生課題音声に着目して発話特徴分析を行い、発話時間に関する特徴量を用いることで点数では見出せない認知機能の低下の兆しを見出せると示した。一方、日常的に利用可能な認知症簡易スクリーニングの実現を目的として、人型ロボットとの日常会話音声における近時記憶に関する話題においても、健常者と MCI 患者間の発話特徴に違いが存在するか分析する。

日常会話における近時記憶では、昨日の食事の内容を覚えているかどうかがあり、近時記憶に関する質問から MCI 患者を見出せる可能性がある。そこで、人型ロボットとの日常会話音声において、近時記憶に関する質問である昨日の食事についての質問を含んだ話題に着目し、発話特徴分析を実施した。この食べ物に関する話題は以下のシナリオで進行する。

- 1) 昨日は何を食べましたか？
- 2) その食べ物は美味しかったですか？
- 3) その食べ物ってどのようにして作るのですか？
- 4) 食べ物といえば、私はリンゴが大好きなんです。
- 5) あなたの好きな食べ物を教えてもらえますか？
- 6) その食べ物はどのような所が好きなんですか？

これら人型ロボットからの問いかけや自己開示に対して実験参加者は返答をする。なお、実験参加者の返答に合わせた返事を人型ロボットが行っているが、その返事に対して答える人と答えない人が居たため、その部分は分析対象から外した。

5.2 結果

人型ロボットとの日常会話音声における発話特徴分析の結果を表7に示す. 反応時間長, 発話時間長, Shimmer(local)において有意差を得たと共に, 効果量の基準 0.2 (Small)を上回った. また, 効果量の基準は下回ったものの, 回答時間長において有意傾向が得られた.

5.3 考察

反応時間長は, 健常者に比べ MCI 患者の方が平均的に長くなっており, MCI 患者は人型ロボットの問いかけが終わってから返答し始めるまでに時間を要している. 認知機能低下と共に回答を思考するのに時間を要したため, 反応時間が遅くなったと考えられる. 発話時間長においては, 健常者に比べ MCI 患者の方が平均的に短い傾向がある. MCI 患者は人型ロボットの問いかけに対して健常者に比べ短い返答をしている. 池田[15]は, 認知症患者の多くは寡黙であると指摘しており, MCI 患者も同様に寡黙になる傾向があったため, 発話時間長が短くなったと考えられる. Shimmer(local)において有意差が得られたが, Jitter(local)では有意差が得られず, 文献[10]の際に示した結果とは異なる結果となった. 今回は人型ロボットとの日常会話の中でも, 近時記憶に関する質問を含んだ食事の話題に絞り分析を行った. 対象とした話題の違いで違いが出た可能性がある. 今後, 近時記憶に関する質問を含まない話題についても分析を行い, どのような話題が認知機能低下の兆しを見出すのに有効か検討を行う必要がある. また, 食事の話題は頻繁に食事を作っていると考えられる女性の方が具体的に食べた食事の内容や作り方を答えていた. 性別によって差が出る可能性があるため, 今後, 性別を分けての分析を行う予定である.

遅延再生課題音声の分析結果(表4)と比較すると, 回答時間長は遅延再生課題音声で有意差が得られ, 人型ロボットとの日常会話音声では有意傾向であった. 回答時間長は1ターン中の実験参加者の発話開始点から終了点までの時間であり, 発話時間長, 無音時間長, フィラー長などの合算値である. このため, 遅延再生課題音声では無音時間

長, フィラー長の2つの特徴量で有意差が得られたため回答時間長で有意差が得られたが, 人型ロボットとの日常会話音声では発話時間長のみで有意差が得られたため, 有意傾向になったと考えられる. 反応時間長では, 遅延再生課題音声, 人型ロボットとの日常会話音声共に, 健常者に比べ MCI 患者の方が長くなる傾向が得られたが, 有意差は人型ロボットとの日常会話音声のみであった. 遅延再生課題は3つの単語という回答が決まっており, どれか1つの単語だけでも思い出せば回答が可能なのに対し, 日常会話は回答が決まっていない. MCI 患者は認知機能低下によって回答を思考するのに時間を要したため, 反応時間長は人型ロボットとの日常会話音声のみで有意差が出たのであろう. 反応時間長は人型ロボットとの日常会話特有の特徴量である可能性がある. 発話時間長は, 人型ロボットとの日常会話音声で有意差が得られたが, 遅延再生課題音声において有意差は得られなかった. 有意差が人型ロボットとの日常会話音声のみであったのは, 遅延再生課題の発話内容が基本的に離散的な単語のみであるのに対して, 人型ロボットは日常会話であり, 回答は単語だけの発話に限らないという日常会話の性質が影響したと考えられる. また, 遅延再生課題音声では健常者に比べ MCI 患者の発話時間長が長い, 人型ロボットとの日常会話音声においては, MCI 患者の方が短いという逆の傾向が見られた. 遅延再生課題の音声を聞くと, MCI 患者は近時記憶能力の低下により遅延再生課題の単語を回答し終えるまでに何度も同じ単語を繰り返すなどしていた. このため発話時間長が長くなっていた. 人型ロボットとの日常会話音声において発話時間長が短いのは, 前述したように MCI 患者は寡黙になったためだと考えられる. 発話時間長も反応時間長と同様に, 人型ロボットとの日常会話特有の特徴量である可能性がある.

Shimmer(local)は遅延再生課題音声, 人型ロボットとの日常会話音声共に同様の傾向が得られた. 効果量を見ると遅延再生課題では 0.5(Medium)の基準を超えており, 人型ロボットとの日常会話音声の 0.2(Small)に比べ, 遅延再生課題の方が Shimmer(local)に大きな差があると言える. 遅延再生

表6 遅延再生課題における満点取得者の分析結果

特徴量名	健常者	MCI 患者	P value	Cohen's d
	Mean (SD)	Mean (SD)		
回答時間長	3.18 (2.38)	3.91 (2.39)	0.34	0.30
反応時間長	0.33 (0.82)	0.56 (1.20)	0.49	0.22
発話時間長	2.08 (0.69)	2.11 (0.67)	0.89	0.04
無音時間長	1.01 (2.00)	1.62 (1.80)	0.32	0.32
無音時間割合	0.17 (0.25)	0.32 (0.22)	0.049 *	0.64
フィラー長	0.09 (0.29)	0.13 (0.43)	0.74	0.11
フィラー割合	0.02 (0.08)	0.02 (0.05)	0.89	0.04
Jitter (local)	0.03 (0.01)	0.04 (0.01)	0.43	0.25
Shimmer(local)	0.12 (0.03)	0.12 (0.03)	0.44	0.25

・ p < .1 * p < .05 ** p < .01 *** p < .001

表7 人型ロボットの日常会話音声の分析結果

特徴量名	健常者	MCI 患者	P value	Cohen's d
	Mean (SD)	Mean (SD)		
回答時間長	9.06 (12.30)	7.34 (7.90)	0.06	0.17
反応時間長	1.07 (1.08)	1.41 (1.56)	<0.01 ***	0.25
発話時間長	6.36 (8.58)	4.64 (4.73)	<0.01 ***	0.25
無音時間長	2.54 (3.75)	2.27 (3.16)	0.57	0.05
無音時間割合	0.18 (0.18)	0.20 (0.20)	0.15	0.13
フィラー長	0.46 (0.90)	0.37 (0.92)	0.26	0.10
フィラー割合	0.05 (0.12)	0.04 (0.12)	0.61	0.05
Jitter (local)	0.03 (0.01)	0.03 (0.01)	0.11	0.15
Shimmer(local)	0.12 (0.03)	0.12 (0.03)	0.02 *	0.23

・ p < .1 * p < .05 ** p < .01 *** p < .001

課題音声は認知機能検査音声のため回答には単語を思い出す必要があるのに対して、日常会話音声の食事の話題は、朝、昼、夜、どれかの食事内容を思い出せば回答することが可能である。遅延再生課題に比べて人型ロボットとの日常会話は実験参加者への負荷が少ないと考えられ、遅延再生課題と日常会話の性質の違いが影響した可能性がある。また、遅延再生課題において有意差を得られたが、人型ロボットとの日常会話音声において有意差が得られなかった無音時間長、無音時間割合、フィラー長も、決まった回答となる遅延再生課題と決まった回答の無い人型ロボットとの日常会話という性質の違いが影響したと考えられる。

6. まとめ

本研究では、近時記憶に関する質問を用いて MCI 患者発見の可能性を検討するために、近時記憶を評価する遅延再生課題提示時の収録音声から 9 種類の特徴量を抽出し、健常者と MCI 患者間の発話特徴分析を行った。人型ロボットとの会話において近時記憶を必要とする場合に、健常者と MCI 患者の発話特徴に差異があるか分析した。結果、遅延再生課題音声において、調査した特徴量 9 種類のうち 6 種類で有意差と有意傾向が得られた。近時記憶に関する課題において健常者と MCI 患者の発話特徴が異なることを示した。一方、人型ロボットとの日常会話では、遅延再生課題音声とは有意差の出た特徴量や得られた傾向が異なるものの 9 種類中 4 種類の特徴量において有意差と有意傾向が得られた。近時記憶に関する質問や話題に着目することで、MCI 患者発見に繋げられる可能性を示した。今後は有意差が得られた特徴量を用いて自動分類実験を行い、近時記憶に関する質問から健常者と MCI 患者の判別可能か検討を行う予定である。

謝辞

本研究は JSPS 科研費 JP19H01084 の助成を受けたものである。

参考文献

- [1] Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E., "Mild cognitive impairment: clinical characterization and outcome," *Archives of neurology*, 56(3), pp.303-308, 1999.
- [2] Folstein, M. F., Folstein, S. E., & McHugh, P. R. ., "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician," *Journal of psychiatric research*, 12(3), pp.189-198, 1975.
- [3] 鳥羽研二, "1. 認知症高齢者の早期発見 臨床的観点から," *日本老年医学会雑誌*, Vol.44, No.3, pp.305-307, 2007.
- [4] Kato, S., Homma, A., Sakuma, T., & Nakamura, M., "Detection of mild Alzheimer's disease and mild cognitive impairment from elderly speech: Binary discrimination using logistic regression," *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 5569-5572, 2015.
- [5] Roark, B., Mitchell, M., Hosom, J. P., Hollingshead, K., & Kaye, J., "Spoken language derived measures for detecting mild cognitive impairment," *IEEE transactions on audio, speech, and language processing*, Vol.19, No.7, pp.2081-2090, 2011.
- [6] Khodabakhsh, A., Yesil, F., Guner, E., & Demiroglu, C., "Evaluation of linguistic and prosodic features for detection of Alzheimer's disease in Turkish conversational speech," *EURASIP Journal on Audio, Speech, and Music Processing*, Vol.2015, No.1, pp.1-15, 2015.
- [7] 外山翔平, 齋藤大輔, 峯松信明, 小林朗子, & 小島敏昭, "音声・言語素性を用いた機械学習による認知症進行度推定に関する予備的検討," *情報処理学会研究報告, SLP-117*, No. 8, pp.1-6, 2017.
- [8] 柴田崇徳, & 和田一義, "アザラン型ロボット「パロ」によるロボット・セラピーの効果の臨床・実証実験について," *日本ロボット学会誌*, Vol.29, No.3, pp.246-249, 2011.
- [9] 二宮 恒樹, "コミュニケーションロボット「PALRO (パルロ)」の紹介とさがみロボット産業特区における取り組み," *日本ロボット学会誌*, Vol.33, No.8, pp.607-610, 2015.
- [10] 吉井謙太, 木村大毅, 小杉晋央, 新川香, 高瀬俊郎, 小林正朋, 山田康智, 根本みゆき, 渡辺亮平, 塚田恵鯉子, 太田深秀, 東晋二, 根本清貴, 新井哲明, 西村雅史, "人型ロボットとの日常会話音声を用いた認知症簡易スクリーニングのための基礎的検討," *情報処理学会研究報告, SLP-133*, No.7, pp.1-4, 2020.
- [11] ソフトバンクロボティクス株式会社, "Pepper", <https://www.softbank.jp/robot/pepper/>, (取得日: 2020/10/26).
- [12] 西村雅史, 小林悠一, 桐山伸也, 峰野博史, "生体音と環境音の同時収録による高齢者の行動および身体状態認識に関する検討," *音響学会講演論文集*, 2-4-9, pp. 1309-1310, 2015.
- [13] P. Boersma, D. Weenink., "Praat, a system for doing phonetics by computer," *Glott International*, Vol.5, No.9, pp.341-345, 2001.
- [14] 水本薫, 武内理., "研究論文における効果量の報告のために。基礎的概念と注意点," *英語教育研究* 31, pp. 57-66, 2008.
- [15] 池田学., "認知症者のコミュニケーション," *高次脳機能研究 (旧 失語症研究)*, 35.3, pp. 292-296, 2015.