

音響解析による認知機能低下の検知

Automatic Detection of Cognitive Impairments through the Acoustic Analysis of Speech

張 亜明
Yaming Zhang

小川 祐輝
Yuki Ogawa

南雲 亮佑
Ryosuke Nagumo

細川 満春
Mitsuharu Hosokawa

岡田 崇志
Takashi Okada

笹部 孝司
Kohji Sasabe

要 旨

高齢化社会の到来に伴い、認知症およびその前段階である軽度認知障害の早期検知に対して社会的な注目が集まっている。従来の検知手法は性能が高い反面、検知時間が長い、検査員の資格保有が問われるなどの課題がある。そこで、これらの課題を解決するために、1万人超の音声データを取得し、短時間の簡単な発話タスクを新たに開発し、発話の流暢（りゅうちょう）性を反映する特徴量から認知機能の低下を推定する機械学習モデルの構築を行った。結果、認知機能低下を高い精度（AUC = 0.795）で推定可能であることが示唆された。

Abstract

Along with the rapid growth and aging of population in Japan, early detection of dementia and mild cognitive impairment has been receiving more and more attention. The existing screening tools for cognitive impairments are usually time-consuming, and also need to be administered by trained professionals. In order to solve these issues, we built a machine learning model using the acoustic features extracted from the speech of more than 10000 elderly individuals. The results reveal that the individuals' acoustic features reflecting fluency can be employed to discriminate between healthy controls and those with cognitive impairments, with high accuracy (AUC = 0.795).

1. はじめに

内閣府の報告によると、2015年時点の日本在住の65歳以上の人口のおよそ15%が認知症患者である[1]。今後の社会人口構成の高齢化率上昇による認知症患者数の増加は、社会保険制度の維持と高齢者や家族・介護者などの生活の質の向上などを鑑みると、大きな社会課題である。近年、認知症の前段階である軽度認知障害(Mild Cognitive Impairment)において、認知機能低下の早期発見とそれに伴う運動などの適切な介入を施すことで、認知機能の改善、もしくは認知機能低下の抑制が期待できることが注目されている[2]。

2. 認知機能低下の早期発見

認知機能低下の早期発見を目的とした検査として、Mini-Mental State Examination (MMSE) [3]やNational Center for Geriatrics and Gerontology-Functional Assessment Tool (NCGG-FAT) [4]などの神経心理学的検査がよく使われている。MMSEは、被験者の認知機能状態を幅広く測る検査方法であり、NCGG-FATは、記憶機能、注意機能、実行機能、情報処理機能の4つの認知機能の低下を網羅的に調べる検査であり、その検査結果から中核症状に応じたケアサービスを展開することができる。一方、これらの検査は、「検査時間が長い」、「検査員の資格保有が問われる」という課題もあり、認知機能低下を早期発見する

うえで阻害要因となっている。そこで、より迅速かつ簡便な検査が求められている。

健忘型MCIやアルツハイマー型認知症 (Alzheimer's Disease) の初期段階において、エピソード記憶、意味記憶などの記憶機能に障害が現れるが、最近では、認知機能の低下に伴って発話においても障害が現れることが報告されている[5][6]。これらには、発話の流暢性を反映する発話速度の低下、ポーズ割合の増加といった音響的特徴が見られる。そして、これらの音響的特徴を利用して認知機能の低下を予測した報告もある[7][8]。

Königらは、神経心理学的検査において用いられる計算、情景画叙述、復唱、カテゴリー語想起のタスクから音声を取得し、記憶機能を測ることで、認知機能低下を予測している[8]。この方法では、従来の神経心理学的検査で用いるタスクを採用することで、神経心理学的考察ができる利点があるが、タスク応答により過大なストレス負荷を被験者に与えるという欠点もある。

そこで筆者らは、迅速かつ簡便な検査であり、また先行研究に着目して、記憶機能に頼らずに被験者に低負荷の音声検査タスクを新たに考案し、軽度認知障害(MCI)、認知症(AD)を予測する機械学習モデルを構築し、認知機能低下を推定する可能性を検証した。MCI、ADの詳細な定義は、3.2認知機能分類にて後述する。

3. 音声による認知機能予測

本研究は、日常生活の行動から認知機能低下の早期の兆候を検知する技術開発の一環であり、日常生活の活動、歩行を長期間モニタリングする検知システムの開発と並行して進めている[9].

音声を用いて認知機能の低下を検知する先行研究の多くは、被験者数が100人未満のものであるが、筆者らは国立研究開発法人国立長寿医療研究センターと共同研究を行い、2017年9月から2018年7月にかけて愛知県東海市と豊明市において大規模コホート調査を行い、計1万人以上の被験者である高齢者の脳と体の健康状態を測定した。認知機能に関する測定項目は、神経心理学に基づく検査（MMSE, NCGG-FAT）などだけでなく、被験者の音声検査を含んでいる。被験者個人が特定できないように、取得した認知機能検査データと音声データを識別番号で管理し分析を行った。なお、本研究は、共同研究先国立長寿医療研究センターの倫理審査の承認のもと行った。

詳細なる検査手法や分析手法について、以下に述べる。

3.1 被験者

愛知県東海市、豊明市に住む65歳以上の高齢者10041人が本研究に参加した。認知機能による分類を行う前に、全被験者のなかから分析除外項目に相当する被験者を排除した。分析除外項目は以下のとおりである：既往歴（パーキンソン病、脳卒中、鬱、認知症など）、身体機能低下が見られる（日常的に食事、整容、歩行、入浴、階段昇降ができない）、または以上の項目に関する情報がないもの。

分析除外要件を考慮した結果、全被験者10041人のうち、分析対象者は8881人（男3959、女4922、年齢層65~96歳、平均年齢74.3±5.7歳）となった。

3.2 認知機能分類

認知機能検査の内容について述べる。今回、MMSEとNCGG-FATを認知機能検査手法として採用した。MMSEはADの分類に[10]、またNCGG-FATはMCIの分類に[11]、高い予測性能を示すことが知られている。

MMSEは時間の見当識、場所の見当識、即時再生、計算、遅延再生、物品呼称、文の復唱、口頭指示、書字指示、自発書字、図形模写の11項目のタスクからなる30点満点の認知機能検査であり、一般的な認知機能の状態を調べることが可能である。

NCGG-FATは10 Words Test（10個の単語の即時・遅延再生・再認）、Trail Making Test（数字を順番に選ぶテスト、TMT）Part A、TMT Part B および Symbol Digit Substitution

Test（符号と数字の置換）の4項目のタスクからなる認知機能検査であり、記憶機能、注意機能、実行機能、情報処理機能の認知機能の状態を調べる。その成績を、教育歴、年齢、性別を考慮したうえで平均値から算出し、 1.5σ （ σ ：標準偏差）以上離れている場合は1点を加算する。

本研究では、この2種類の認知機能検査手法を用い、認知機能低下の深刻度などの組み合わせにより便宜上、以下の3群に分類した。分類の基準値を第1表に示す。第2表にこの3分類の被験者の分布状況を示す。各群（健常者、MCI、AD）の分析対象者数は、それぞれ6343（71.4%）、1601（18.0%）、937（10.6%）となった。また、NCGG-FATが調べる4つの認知機能において低下が見られる人数も含まれている。同時に複数の機能低下が見られる被験者がいるため、4群の人数の合計が総被験者数を超えている。

第1表 MMSEとNCGG-FATの点数による認知機能分類

Table 1 Classification of participants based on MMSE and NCGG-FAT scores

		MMSE	
		0~23	24~30
NCGG-FAT	0	AD	健常者
	1~4		MCI

第2表 認知機能分類後の被験者分布状況

Table 2 Characteristics of participants in the three cognitive groups: healthy controls, MCI, and AD

	健常者 (n=6343)	MCI (n=1601)	AD (n=937)
被験者情報の分布			
年齢 [歳]	73.5 ± 5.4	74.9 ± 5.6*	78.1 ± 5.9*
教育歴 [年]	12.0 ± 2.4	11.6 ± 2.4*	10.6 ± 2.2*
男性 [人 (%)]	2658 (41.9)	757 (47.3)	544 (58.1)
認知機能			
MMSE [点]	27.7 ± 0.9	26.8 ± 1.9	21.7 ± 1.9
NCGG-FATによる機能低下 [人 (%)]			
記憶機能	-	548 (34.2)	274 (29.2)
注意機能	-	474 (29.6)	161 (17.2)
実行機能	-	677 (57.7)	425 (45.3)
情報処理機能	-	160 (10.0)	159 (17.0)

*：健常者と比較して有意差が見られた($p < 0.01$)

3.3 音声検査タスク

今回実施された音声検査は以下4つのタスクからなっている。閉鎖子音をたくさん含むことによって、脳内で調音運動を制御する能力が調べられる。従来の神経心理学検査は20分以上の時間がかかるのに対して、本タスクは約5分で完成できる。

a) 母音発音タスク：日本語の五母音/a/, /i/, /u/, /e/, /o/をそれぞれ3秒間連続して発音する。

- b) 早口言葉タスク: 閉鎖子音 ($/p/, /t/, /k/$) を多く含む文章「きたからきたかたたたき」をできるだけ早く3回繰り返して読み上げる。
- c) 交互変換運動タスク: 閉鎖子音 ($/p/, /t/, /k/$) から始まる単音節語「ぱ」, 「た」, 「か」をできるだけ早く5秒間連続して発音する。
- d) 文章朗読タスク: 以下の指定された文章を, 被験者の日常会話における通常の話速で3回繰り返して読み上げる。

文章1: 「探検家は冒険が大好きです。」

文章2: 「北風と太陽が出ています。」

すべてのタスクでは, 発話内容はパソコンの画面上に提示されているため, 記憶機能はあまり使われない。検査員の指導のもと, 被験者の音声検査は防音材で囲まれるブースにて実施した。録音に使用されるマイク (Zoom^(注1) SGH-6 Shotgun Mic Capsule) はレコーダー (Zoom H6 Handy Recorder) を経由してノートパソコンに接続され, 被験者の口から10 cmを離れた場所に設置されている。サンプリング周波数は44 kHzに設定した。

3.4 特徴量

収録された音声は, パワー, 基本周波数 (F0), フォルマント周波数 (共振周波数 F1, F2) の時系列データにそれぞれ変換した。先行研究において, 認知機能が低下すると, 発話の流暢性に影響が出てくることが知られている[5][6]。そのため, 本研究では主に発話の流暢性を示す時間的特徴を計算した。母音発音タスクでは, 発音の明瞭度を測るため, 周波数領域の情報を抽出した。

a) 母音発音タスク

このタスクでは, 被験者が母音を連続して発音することであり, 正しい調音位置に舌を維持することが必要となるため, 話者の発音明瞭度を反映できる。分析として, パワー, 基本周波数 (F0), フォルマント周波数 (F1, F2) の標準偏差などの特徴量を計算し, フォルマント空間における面積を抽出した。

b) 早口言葉タスク

早口言葉は被験者の実行機能を反映しており, うまく言えない場合には被験者の認知機能が低下している可能性がある[12]。分析として, 課題遂行時間, 発声の回数, 1回当たりの発声時間の平均値と標準偏差を計算した。

c) 交互変換運動タスク

このタスクでは, 被験者が閉鎖子音 ($/p/, /t/, /k/$) から始まる単音節語を発音することであり, 脳内で調音運動

を制御する能力を調べるためによく使われるタスクである。いくつかの先行研究では, このタスクを用い, 健常者からMCIを検出した[13][14]。分析として, 早口言葉タスクと同様の課題遂行時間, 発声の回数, 1回当たりの発声時間の平均値と標準偏差に加え, 発話速度の変動を計算した。

d) 文章朗読タスク

先行研究によると, 認知機能の低下に伴い, 文章朗読において発話テンポに変化が現れ, 具体的には, 調音速度の低下やポーズ割合の増加などの音響的特徴が見られる[7]。分析として, 発話のテンポを反映する発声時間, 発声回数, 発声時間の平均値と標準偏差などの特徴量を計算した。

3.5 機械学習モデル

前節の計算処理によって, 分析対象者の特徴量の計算値が並べられたテーブル (表型) データが得られる。これを機械学習モデルで学習し, その予測性能を評価した。今回は分類問題であるため, 二項ロジスティック回帰 (L2正則) を採用した[15]。その予測性能はROC-AUC (Receiver Operation Curve - Area Under the Curve) を指標として評価している。目的変数に関して, 認知機能低下群 (MCI, AD) またはNCGG-FATによる認知機能低下群 (記憶機能, 注意機能, 実行機能, 情報処理機能) を陽性にし, 健常者群を陰性にした。説明変数に関して, 前節で述べたすべての特徴量が使用された。そのワークフローを以下に示す。

① 予測対象のターゲットラベルを選択する。

② 交差検証は二重ループにし, 内側ループではモデルのハイパーパラメータのチューニングを, 外側のループでは3回の交差検証を10個の異なるランダム定数で行う。

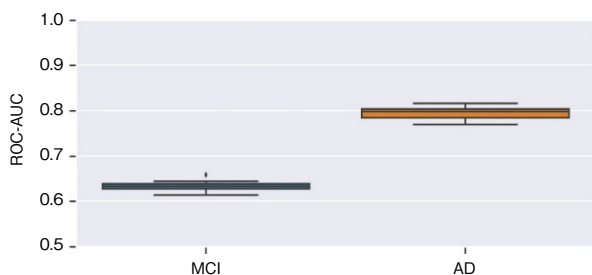
③ 計30回の性能評価の評価値を計算する。そして, それぞれのターゲットに対する評価値の平均値と標準偏差を出し, 予測性能を評価する。

4. 結果

4.1 認知機能低下群の分類性能評価

各認知機能低下群に対する分類性能を比較するために, MCI, ADをそれぞれ健常者と比較した時の分類性能を第1図に示す。予測性能の評価値は, MCI: 0.6341 (± 0.0097), AD: 0.7954 (± 0.0130)である。また, 評価値の間に有意差が見られた ($p < 0.001$)。認知機能低下の深さが大きくなるほど, 分類性能は大きくなるのがわかる。

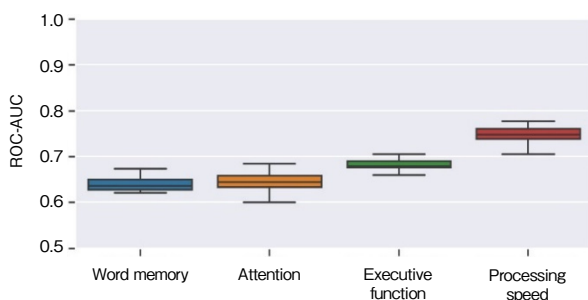
(注1) (株) ズームの日本およびその他の国における登録商標または商標。



第1図 ROC-AUCによる認知機能低下群の分類性能評価
Fig. 1 ROC-AUC score of the cognitive impairment classification tasks to differentiate MCI and AD from healthy controls

4.2 NCGG-FATで定義される各認知機能低下群の分類性能評価

認知機能の低下と音響的特徴の潜在的関連性を探るために、NCGG-FATで検査する記憶機能 (Word Memory)、注意機能 (Attention)、実行機能 (Executive Function)、情報処理機能 (Processing Speed) の4つの認知機能低下群と健常者群とのそれぞれの分類性能を第2図に示す。分類性能は、記憶機能:0.6409 (±0.0158)、注意機能:0.6460 (±0.0194)、実行機能:0.6829 (±0.0120)、情報処理機能:0.7479 (±0.0168)となった。統計分析を行った結果、記憶機能と注意機能の間に有意差は見られなかったが ($p = 0.269$)、実行機能と記憶・注意機能の間、また情報処理機能と他の群3機能の間には有意差が見られた ($p < 0.001$)。



第2図 NCGG-FATで定義される各認知機能低下群の分類性能
Fig. 2 ROC-AUC scores of the impaired individuals in four cognitive domains, that is, word memory, attention, executive function, and processing speed, differentiated from healthy controls

4.3 タスク別の分類性能評価

認知機能低下群 (MCI, AD) と健常者群とのタスク別の分類性能を第3表に示す。

第3表 認知機能低下群と健常者群とのタスク別の分類性能

Table 3 ROC-AUC values for the prediction of cognitive impairments using acoustic features derived from single tasks

タスク	MCI	AD
母音発音	0.5617±0.0122	0.6478±0.0128*
早口言葉	0.6000±0.0093	0.6953±0.0164*
交互変換運動	0.6197±0.0092	0.6950±0.0182*
文章朗読	0.5825±0.0132	0.7018±0.0117*

* : 各群(MCI, AD)の評価値の間に有意差が見られた ($p < 0.001$)

5. 考察

実験の結果より、発話に含まれる音響的特徴から認知機能の低下を推定できることがわかった。また、認知機能低下群の予測性能をSwets[16]が提唱したAUCの評価基準で評価すると、AD (AUC=0.7954) については妥当な性能を示したが、MCI (AUC=0.6341) は比較的に低かった。このことから、認知機能が低いほど予測性能は高くなり、言語能力の低下は認知機能の低下とともに顕著に表れる傾向が示された。先行研究でも、言語の流暢性は認知機能の低下とともに深刻化すると報告されているが [17]、本研究の結果でもこれらの報告と傾向は一致した。

NCGG-FATで定義される4つの認知機能の低下群のそれぞれの予測性能を評価すると、実行機能と情報処理機能の低下群を高い精度で予測することができ、筆者らの音声検査タスクは実行機能と情報処理機能の低下に関して検知感度が高いことが明らかになった。特に早口言葉と交互変換運動のタスクでは、閉鎖子音 (/p/, /t/, /k/) が多く含まれる発音を使用しており、速いスピードで異なる調音位置の間で舌を移動することが必要である。認知機能の低下に伴って、脳内で調音運動を制御する能力が低下するため、発話における調音速度の低下とポーズの増加につながるものと考えられる。

なお、母音発音タスクは他のタスクに比べ、予測性能が低く、認知機能低下の早期発見において、発音の明瞭度はあまり影響されないことがわかった。

過去の音声で認知機能の低下を検知する研究では、記憶機能を使用するタスクで検査するのが一般的である [18]。一方、筆者らの音声検査タスクは定型文の読み上げであるため、従来の認知機能検査で測る記憶機能と注意機能はほとんど使用しない。従って、被験者に対する負担は少なくやさしいにもかかわらず、妥当な予測性能を示すことができた。先行研究と同様に記憶機能を使用するようなタスクを追加すれば、記憶機能と注意機能の低下群に対する予測性能をさらに高められると考える。

6. まとめ

本研究は被験者にとって低負荷な検査の開発を目指し、発話における音響的特徴で認知機能の低下を推定する可能性について検証した。そのため、計1万人以上の音声データを集めた。その結果、従来技術より低負荷で高い予測性能を示すことができた。ただし、ADについては実用に耐え得る十分な高い予測性能と認識しているが、MCIについてはさらなる向上が望まれる。

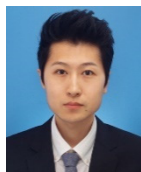
現在、国立研究開発法人国立長寿医療研究センターとの共同研究を継続して行っており、被験者の負荷を極力上げない形で、記憶機能を使用する新たな音声検査タスクを導入することで、MCIの予測性能の改善に挑戦中である。最終的には自由発話の音声からも認知機能低下を高精度で推定できるように改善し、スマートスピーカーなどに実装できることを目指している。

本研究は国立研究開発法人国立長寿医療研究センターと共同で実施を行った内容を基盤としており、関係各位より賜りましたご指導ご助言に感謝致します。

参考文献

- [1] Cabinet Office, "Annual Report on the Aging Society 2017 (Summary)," https://www8.cao.go.jp/kourei/english/annualreport/2017/2017pdf_e.html, 参照 May 7, 2020.
- [2] M. Malek-Ahmadi. "Reversion from mild cognitive impairment to normal cognition: A meta-analysis," *Alzheimers Dis Assoc Disord.* vol. 30, issue 4, pp. 324-330, 2016.
- [3] M. F. Folstein et al., "Mini-mental state. A practical method for grading the cognitive state of patients for the clinician," *J Psychiatr Res.* vol. 12, issue 3, pp. 189-198, 1975.
- [4] H. Makizako et al., "Evaluation of multidimensional neurocognitive function using a tablet personal computer: test-retest reliability and validity in community-dwelling older adults," *Geriatr Gerontol Int.* vol. 13, issue 4, pp. 860-866, 2013.
- [5] R. G. Morris, *The Cognitive Neuropsychology of Alzheimer-Type Dementia*, Oxford, Oxford University Press, 1996.
- [6] N. Biassou et al., "Phonological processing deficits in Alzheimer's disease," *Neurology.* vol. 45, pp. 2165-2169, Dec. 1995.
- [7] F. Martínez-Sánchez et al., "Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects," *Neurología.* vol. 28, no. 6, pp. 325-331, 2013.
- [8] A. König et al. "Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease," *Alzheimer's Dement (Amst).* vol. 1, issue 1, pp. 112-124, 2015.
- [9] K. Abe et al., "Early detection of elderly people with mild cognitive decline," *Panasonic Technical Journal.* vol. 64, no. 2, 2018.
- [10] S. E. O' Bryant et al. "Detecting dementia with the mini-mental state examination in highly educated individuals," *Arch Neurol.* vol. 65, no. 7, pp. 963-967, 2008.
- [11] H. Shimada, H. et al., "Cognitive impairment and disability in older Japanese adults," *PLoS One.* vol. 11, no. 7, 2016
- [12] A. Wutzler et al., "The anticipatory proportion as an indicator of language impairment in early-stage cognitive disorder in the elderly," *Dement Geriatr Cogn Disord.* vol. 36, no. 5-6, pp. 300-309, 2013.
- [13] P. Östberg et al., "Articulatory agility in cognitive decline," *Folia Phoniatrica et Logopaedica.* vol. 61, pp. 269-274, Sept. 2009.
- [14] Y. Watanabe et al. "Oral function as an indexing parameter for mild cognitive impairment in older adults," *Geriatr Gerontol Int.* vol 18, issue 5, pp. 790-798, 2018.
- [15] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, Berlin, 2006.
- [16] J. A. Swets, "Measuring the accuracy of diagnostic systems," *Science.* vol. 240, no. 4857, pp. 1285-1293, 1988.
- [17] M. Alegret et al. "The Role of Verb Fluency in the Detection of Early Cognitive Impairment in Alzheimer's Disease," *J Alzheimers Dis.* vol. 62, no. 2, pp. 611-619, 2018.
- [18] L. Tóth et al. "A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech," *Curr Alzheimer Res.* vol. 15, no. 2, pp. 130-138, 2018.

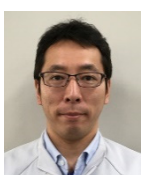
執筆者紹介



張 亜明 Yaming Zhang
ライフソリューションズ社 技術本部
Engineering Div., Life Solutions Company
学術博士



南雲 亮佑 Ryosuke Nagumo
ビジネスイノベーション本部
AIソリューションセンター
AI Solution Center, Business Innovation Div.



岡田 崇志 Takashi Okada
ライフソリューションズ社 技術本部
Engineering Div., Life Solutions Company



小川 祐輝 Yuki Ogawa
ライフソリューションズ社 技術本部
Engineering Div., Life Solutions Company



細川 満春 Mitsuharu Hosokawa
ライフソリューションズ社 技術本部
Engineering Div., Life Solutions Company



笹部 孝司 Kohji Sasabe
ライフソリューションズ社 技術本部
Engineering Div., Life Solutions Company
工学博士