

〈一般研究課題〉 高齢者の対話音声から取得した音響的特徴に基づく
認知症傾向診断

助成研究者 愛知県立大学 入部 百合絵



高齢者の対話音声から取得した音響的特徴に基づく 認知症傾向診断

入部 百合絵
(愛知県立大学)

Detection of early dementia based on acoustic features extracted from dialogue speech of elderly Japanese

Yurie IRIBE
(Aichi Prefectural University)

Abstract :

As Japanese society rapidly ages, the number of people with dementia in the country increases each year. In July 2008, the Ministry of Health, Labor and Welfare launched an "emergency project to raise the quality of medical care and improve daily life for citizens with dementia." Early diagnosis of dementia is listed as one of measures to be taken, so it has been studied through the analysis of physical movement, brain wave measurement, cognitive assessments, etc. However, there have been few studies on the relationship between speech and dementia. For this reason, we approached to automatic detection of dementia based on characteristics of elderly dialogue speech recorded in natural settings. Firstly we conducted a simple diagnostic test used to measure dementia risk with each speaker after recording their speech so that we could analyze possible characteristics of elderly speech which might be linked with dementia. Widely used cognitive assessments used to detect dementia is the revised Hasegawa's Dementia Scale (HDS-R). The test indicates a tendency towards dementia if a subject's score falls below a given threshold. For this study, we adopted the simpler HDS-R. Each subject who provided speech for our corpus was scored using the HDS-R at a later date. As a result, 11 out of the 98 original participants were judged to have a tendency towards dementia. We trained machine learning models on these data to differentiate dementia and non-dementia speech, and detected dementia from dialogue speech based

on the models. As a result, our method obtained overall precision of 88.9%.

1. はじめに

近年、日本では高齢化社会が進み、それに伴い認知症者はさらに増加が予想されている[1]。これらの問題への対策が急務になり、「認知症施策推進総合戦略(新オレンジプラン)～認知症高齢者等にやさしい地域づくりに向けて～」が、2015年1月に厚生労働省より発表された[2]。これは、「認知症の人の意思が尊重され、できる限り住み慣れた地域のよい環境で自分らしく暮らし続けることができる社会の実現を目指し」、認知症者の増加への対策の指針となる戦略として掲げられたものである。認知症の症状は主に記憶障害、アパシー、言語障害、構音障害など、原因疾患や個人によって様々である。こうした認知症に罹患した際の治療法は未だ確立されていないが、早期発見により進行を遅らせることができる[3]。しかし、現在利用されている検査は被験者に対し精神的・肉体的負担を課すものが多い。そこで、本研究では対話音声に着目した。対話音声に含まれる音響情報と言語情報は認知症の症状に表れる特徴を含んでおり、認知症の判別に有用であると予想される。また、現行の方法のように医師や大掛かりな機材なしに収集可能であり、被験者に与えるストレスも少なく検査を行うことができると考えられる。認知症の早期診断を目的とし、対話音声を診断材料とする取り組みは種々挙げられる。田中ら[4]はアバターによる質問機能をもつ認知症検出システムの開発を目的とし、音声特徴、言語特徴、対話特徴等を利用し、認知症の識別モデルを作成している。しかし、この研究では識別モデルの作成のみを着眼点とし、認知症の症状が音声言語特徴にどのように表出されるのかについて分析は行われていない。加藤ら[5]は高齢者の認知機能障害のスクリーニング手法の開発のため、新しい試みとして発話音声の韻律的特徴解析手法及びデータマイニング技術を融合させたアプローチを提案している。115名(平均68.3歳)の高齢者から319の音声資料(認知症傾向あり114, 認知症傾向なし205)を収集し、スペクトル、ピッチ、フォルマントをはじめとした130種の音声韻律特徴を抽出している。この音声韻律特徴を統計的に圧縮し、認知機能検査HDS-R(改訂長谷川式簡易知能評価スケール)との高い相関を示すことで新しい認知機能障害スクリーニング手法としての応用可能性を示唆している。しかし、田中らの研究と同様に、認知症傾向を持つ高齢者の音声の特性についての分析は行われていない。また被験者の平均年齢が比較的lowく、今後の高齢化社会を踏まえるとさらに高い年代の高齢者を中心としたデータを用いる必要がある。

以上のように、対話音声に着目した研究は存在するが、音響情報・言語情報両者に着目したものは絶対数が少なく、知見も十分とは言えない。以上の背景を踏まえ、本研究では高齢者の対話音声を収集し、その言語的・音響的特徴から認知症傾向の判別をおこなうことを目的とする。そのため、高齢者施設において、特に80歳以上の高齢者を中心に対話音声の収録をおこなった。収集したデータから抽出する特徴量は先行研究において有効であるとされる特徴、および認知症の症状として現れる特徴を数値化したものを選択した。その結果、検定により有意差の認められた特徴のみを用いて識別を行うことで、認知症傾向の判別に有用であるかを検証した。

2. 高齢者音声の収録

2.1 概要

2014年5月から2015年2月に亘って高齢者の対話音声进行録音した。収録の様子を図1に示す。以降、収集した音声データについて詳述する。

本研究では、60歳以上的高齢者98人(平均年齢77.2歳)から質問者との対話音声を収録した。話者の年齢や性別の分布を表1に示す。音声収集は、名古屋市内にある大学と4つの高齢者施設で行った。各施設における収録人数と平均年齢を表2に示す。本研究では80歳以上の被験者の割合を高くすることに注力した。被験者98人からは、HDS-R(改訂長谷川式簡易知能評価スケール)の答案とHDS-R実施時の対話音声を収集した。HDS-R(改訂長谷川式簡易知能評価スケール)は日本で最も広く利用されている認知機能検査であり、わが国における認知症のスクリーニングテストとして最も古い歴史があったHDSの質問項目を、現代に則した項目に再検討したものである。年齢や日時の見当識、言葉の記銘などの9項目を口頭で質問し、10分前後で検査を終えることができる。先行研究と同様に本研究でも認知症判断にHDS-Rを用い、発話の収録には卓上型マイクを使用した。収集したHDS-Rをもとに被験者の認知症傾向を求めたところ、98人中11人(11.2%)の被験者が認知症傾向にあると判定された(表1)。

収録した対話音声を人手によって書き起こし文に改め、言語的特徴の分析に用いた。また、認知症傾向ありの11名のうち2名は一部の音声不明瞭であったため音素セグメントの取得が困難で

表1 認知症傾向有無別の年齢と性別の内訳

年齢	認知症傾向なし		認知症傾向あり		合計
	男性	女性	男性	女性	
60-69	4	13	0	0	17
70-79	9	30	0	2	41
80-89	6	21	1	3	31
90-99	3	1	0	5	9
合計	22	65	1	10	98

表2 収録人数と平均年齢

収録場所	収録人数	平均年齢
高齢者施設 A	11	85.1
高齢者施設 B	11	83.5
高齢者施設 C	17	80.6
高齢者施設 D	19	81.4
名古屋大学	44	70.8



図1 施設での音声収録の様子

あった。そのため、認知症傾向ありの9名とほぼ同数の認知症傾向無し13名の音声データに対し、高齢者音声コーパスS-JNAS[6]により学習した音響モデルを用いた強制アライメントを実施することで音素アライメント情報を取得した。対話音声は約1時間2分に及ぶ22名の音声である。音素アライメントは音響的特徴の分析に利用した。被験者を認知症傾向の有無により2グループに分け、言語および音響の両面から認知症傾向有無に対する比較分析を実施した。

2.2 音声収録の手順と対話音声

収録を開始するにあたって、話者に収録の目的やその用途について説明し、話者全員から同意書を得るなどインフォームドコンセントを実施した。録音開始前には話者に収録の手順を説明するとともに、読み上げの練習時間を設けた。読み上げ文は漢字かな混じりのテキストであるが、円滑に文章を読み上げることのできるように、すべての漢字にかな文字のルビを付けたテキストを話者に渡した。各人が読み上げる文数は50文であるが、話者によっては複数の文を休憩なしで読むことが困難だったため、話者の体調に応じて途中休憩を取ることにした。HDS-Rによる認知テストは読み上げ音声の録音後に実施し、インタビュアーが話者に直接問いかけながら回答してもらった。本音声データは全ての収録場所において同じ録音装置を用いて収録し、収録用マイクには卓上タイプのAudio-Technica AT9930を、レコーダーにはTASCAM社DR-05VERSION2を用いた。

3. 収集した音声データの分析

本研究では、98名の被験者から認知症傾向を測るためのテストHDS-Rを実施し、その際に得た会話音声およびHDS-Rのテスト結果を収集した。HDS-Rテストの結果をもとに認知症傾向を判定し、認知症傾向をもつ場合ともたない場合の音響的・言語的特徴の違いを分析する。分析する特徴については以降に詳述する。

3.1 分析に用いた言語的特徴量

HDS-R実施時の会話音声の書き起こし文を形態素解析器MeCab [7]によって形態素の情報を付与し、以下に詳述する指標によって分析を行った。なお、分析するデータの総形態素数は12,866であった。以下に実際に対話音声から抽出した特徴量について記載する。

- 言語的特徴
 - 語彙
 - ・ 固有名詞割合／一般名詞割合
 - ・ TTR /ユールのK 特性値
 - ・ 代名詞割合
 - ・ 語彙レベル
 - 言い淀み、迂言
 - ・ フィラー割合
 - ・ 異なり名詞割合
 - 各品詞の割合(10品詞)

【固有名詞割合／一般名詞割合】 認知症の症状の一つとして、日常で使う名詞が理解できなくなるというものがある。例えば、目の前に電話を置かれ、その名称を問われても答えることができず正

解を聞いても納得できなかつたり、家族の名前を忘れてしまうという例が挙げられる[8]。この言語流暢性に関わる症状は初期にも現れることがあり[9]、認知症のスクリーニングのための判断材料として重要である。そのため、これらの症状を語彙量の豊富さや用いる語彙の難度を定量化することができないか検証した。前述したように、認知症の症状にはよく知った人の名前(固有名詞)、日常で使う物の名前(一般名詞)などがわからなくなる、あるいは言葉として出てこなくなるというものがある。そこで、対話中に含まれる固有名詞、一般名詞の割合が認知症傾向の有無の判別指標として利用できないかを検証した。書き起こし文を形態素解析器MeCabによって解析し、品詞の付与および分かち書きを行う。MeCabの定義する品詞に従い固有名詞、一般名詞を計上することで、全体の各名詞の割合を算出する。

【代名詞割合】 認知症に罹患すると代名詞を多く用いるようになることは、症状の一つとして明らかになっている[8]。観察法の認知症重症度評価であるSSDS(Screening Scale for Dementia Severity)においても境界兆候のひとつとして扱われ[10]、会話音声から認知症の検出を試みる際、有用な情報となることが期待される。発話中に含まれる代名詞の割合を数値化することで、短い対話中であっても同様の傾向が現れるかを検証する。

【語彙レベル】 用いた名詞の難易度を測り、各被験者の有する語彙レベルとする。難易度は、日本語教育語彙表[11]に基づく。日本語教育語彙表は、各語彙を日本語教育上のレベルで6段階に分け、語種や品詞などと共に記載したものである。この内、下位2段階を除いた中級前半、中級後半、上級前半、上級後半に当たる語のみを計上した。

【異なり名詞割合】 認知症の症状は人や原因によって様々だが、その中の一つに、特定の語彙を多用する語彙収束傾向がある[8]。また、直前と同じ話を繰り返すことも症状の一つとされており、これらの特徴を現すために定義するものである。

【各品詞の割合】 先行研究のなかには、名詞や代名詞、既存の語彙量などの特徴について触れているものもある。しかし、その多くは名詞のような体言のみを対象としたものであり、その他の内容語や機能語に触れているものは少ない。それは、観察法や対話を行う中で認知症を罹患した人とそうでない人との相違点を感じ取る際、最も強く意味を持つ名詞に注目が集まりやすいため、知見が豊富であることが原因と考えられる。そこで、本研究では形態素解析器MeCabで解析可能な、主だった品詞すべてに対して割合の算出を行い、認知症傾向との関連を検証する。対象となるのは、連体詞、接頭詞、動詞、形容詞、副詞、接続詞、助詞、助動詞、感動詞、名詞である。

3.2 分析に用いた言語的特徴量

被験者98名の中でHDS-Rの採点結果において認知症傾向ありと判断された11名のうち、2名は音素セグメントの取得が困難であった。そのため、認知症傾向ありの9名とほぼ同数の認知症傾向なし13名の音声データに対し、高齢者音声コーパスS-JNASにより学習した音響モデルを用いた強制アライメントを行い、約1時間2分に及ぶ22名の対話音声の音素アライメント情報を取得した。この音素アライメント情報および収集したHDS-R実施時の対話音声から、音響的特徴の分析を行った。認知症の罹患によって調音に影響が生じることが知られているものの、具体的にどのような特徴として現れるのかは明らかになっていない。そこで、本研究では認知症傾向の有無にかかわらず、高齢に伴って生じる調音の特徴に関する先行研究[12]をもとに音響的特徴を選択した。

- 音響的特徴
 - 無音区間
 - ・挿入位置の傾向
 - ピッチ
 - ・子音(C) から母音(V) へのわたりの部分の傾き
 - ・子音の立ち上がり部分の傾き
 - パワー
 - ・子音(C) から母音(V) へのわたりの部分の傾き／高低差
 - ・子音の立ち上がり部分の傾き
 - フォルマント周波(Hz)
 - ・母音フォルマント
 - ・子音フォルマント
 - ・母音の五角形の面積

【無音区間】 音声は実際に音声が発せられている音声区間と、音素の前後に生じる無音区間で構成される。本稿では、無音区間の挿入される位置と、無音区間の占める時間(s)に着目して分析を行う。無音区間、音声区間の弁別は対話音声に付与されたアライメント情報に基づいて行い、データの始まりや終わり位置に含まれる無音区間は除去の上算出する。無音区間が挿入される位置の傾向の分析は、子音の立ち上がり前と子音(C)・母音(V)間について行う。子音の立ち上がり前の無音区間ではその後続く子音、CV 間の無音区間ではその前に発せられた子音ごとに分けて検証する。無音区間の時間についても、無音区間が挿入される位置の傾向の分析と同様に区別し、その上で無音区間の時間について検証する。

【ピッチ】 ピッチは基本周波数とも呼ばれる声の高さを表すパラメータである。ピッチの抽出法は多数提案されているが、本稿ではピッチをRAPT(A Robust Algorithm for Pitch Tracking)により求める。子音(C) から母音(V) へのわたりの部分の微分値では、子音と母音の切り替わる部分の前後3 フレームにおけるピッチを抽出し、その3 フレーム間での微分値を算出する。なお、本稿では1 フレームを20ms とする。今回は子音についてのみ区別し、続く母音の種類の違いはおこなっていない。子音の立ち上がり部分の微分値では、全ての子音の立ち上がり部分3 フレームでのピッチ(Hz)を抽出し、その3 フレーム間での微分値を算出する。1 フレームは前述したように20ms とする。認知症の症状の一つとして発話開始困難が現れることが明らかになっている[13]。その”発話開始困難”が現れると仮定し、抽出をおこなう。

【パワー】 瞬時音声パワーを抽出し、分析をおこなう。本稿では、パワーを $10 \log(\text{振幅})^2$ とし、認知症傾向の有無での変化について考察する。まず子音(C) から母音(V) へのわたりの部分の微分値／高低差について、ピッチと同様に1 フレーム20ms とし、前後3 フレーム間での微分値および高低差を抽出する。ただし、高低差は(パワーの最大値-最小値)により求める。今回は子音についてのみ区別し、続く母音の区別はおこなっていない。次に、子音の立ち上がり部分の微分値／高低差についてもCV のわたりの部分と同様に分析を行い、発話開始困難が現れる子音を探す。

表3 有意差が認められた言語的特徴

特徴	認知症傾向あり		認知症傾向なし		t検定
	平均	標準偏差	平均	標準偏差	t値
代名詞割合	3.06×10^{-2}	1.24×10^{-2}	1.86×10^{-2}	0.84×10^{-2}	2.54*
語彙レベル	0.367	0.076	0.446	0.079	2.69**
フィラー割合	1.79×10^{-2}	2.31×10^{-2}	3.70×10^{-2}	3.20×10^{-2}	2.17*
異なり名詞割合	0.452	0.096	0.568	0.114	3.02**
副詞割合	2.25	1.50	3.48	1.36	2.21*
感動詞割合	4.98	3.37	7.86	4.60	2.24*
名詞割合	25.73	2.36	23.72	3.90	2.19*

* $p < 0.05$, ** $p < 0.01$

表4 有意差が認められた音響的特徴(一部のみ記載)

特徴	認知症傾向あり		認知症傾向なし		t検定
	平均	標準偏差	平均	標準偏差	t値
子音/N/の立ち上がり部分のパワー微分値 (標準偏差)	3.55	0.78	5.45	1.56	3.52*
子音/s/のわたり部分のパワー高低差 (標準偏差)	5.16	0.72	4.00	0.81	3.29**
子音/z/の第二フォルマント	1991.95	73.60	1704.91	228.35	2.95*

** $p < 0.01$

【フォルマント周波数(Hz)】声道が作り出す共振特性の各共振周波数であるフォルマント周波数を抽出し、分析をおこなう。母音のフォルマントは、第1, 第2 フォルマントを、子音のフォルマントは第1, 第2, 第3 フォルマントを抽出し、認知症傾向との関連を検証する。フォルマントの抽出法はいくつか提案されているが、本稿では線形予測分析(Linear Predictive Coding)に基づくフォルマント周波数の抽出をおこなう。母音の五角形面積では縦軸を第1 フォルマント、横軸を第2 フォルマントとして母音のフォルマントを分布した際に形成される五角形の面積を求める。この面積が大きいほど調音がしっかりと行われていると考えられる。

3.3 分析結果

言語的・音響的特徴量を対話音声から抽出後、認知症傾向と相関のある特徴を明らかにするために、認知症傾向を有している被験者から収録した対話音声と有していない被験者の対話音声に分けて、それらの間で有意差が認められた特徴量について述べる。有意差を求める際に利用した検定は両側t検定である。

言語的特徴、音響的特徴において有意差が認められた特徴をまとめる。言語的特徴で有意差が認められたものを表3に、音響的特徴で有意差が認められたものを表4に記す。なお、音響的特徴量に関しては、紙面の関係上、有意水準 $p < 0.01$ で有意差が認められた特徴のみ記載するが、有意水準 $p < 0.05$ も含めると32種類の特徴量に関して有意差が確認された。

代名詞割合、フィラーの割合、異なり名詞割合、語彙レベル、副詞割合、感動詞割合、名詞割合の7種に対して有意差が認められた。代名詞割合は認知症傾向の有無で5%水準の有意差が現れた。本稿では割愛したが、年代ごとの認知症傾向有無別に代名詞割合の結果を確認したところ、年代が

上がると認知症傾向ありの被験者は代名詞割合が増加していた。一方で認知症傾向なしの被験者は年齢が上がると代名詞割合が低下しており、年代が上がると認知症傾向ありの被験者との間に差が大きくなることが分かった。このことから、代名詞割合における有意差は加齢が原因となったものではないことは明らかであり、代名詞割合の増加は、短い対話中でも認知症傾向の判別に有用である可能性がある。また、異なり名詞割合、語彙レベルにおいては1%水準の有意差が認められ、認知症傾向ありの群は対話中に用いる名詞の種類が少なく、語彙の難度も低くなる可能性が示された。副詞／感動詞割合は、認知症傾向のある被験者よりも認知症傾向のない被験者の方が高い値となっている。感動詞は、「おや、これは」、「はい、わたしです」のように感動・呼びかけ・応答などを表す語である。この感動詞の割合が低くなっている点から、認知症の症状の一つである感情の発露の低下や、コミュニケーションへの意欲の低下などが現れているものと推測される。また、副詞は「かなり少ない」、「ゆっくり歩く」などのように後に続く語を修飾する働きを持つ。ほとんどの場合修飾語としてのみ働くため、意図する内容を伝えるために省略することのできる品詞であり、副詞の割合の高さは発話した文章における表現の豊かさであると言い換えることもできる。よって、認知症傾向のある被験者の副詞割合が低いことは、その話す内容の表現の豊かさの低下を示している可能性が考えられる。しかし、同様に修飾語として働く連体詞の割合には有意差が認められておらず、この副詞と連体詞の違いについて今後検証する必要がある。

音響的特徴では無音区間挿入回数、母音・子音のわたりの部分のパワーの高低差、ピッチの変化率、母音・子音のフォルマントなどの特徴を抽出し、言語的特徴と同様に認知症傾向の有無における有意差を求めた。その結果、32種類の音響的特徴量に対して有意差が確認された。

4. 認知症傾向の判別

言語的・音響的特徴の抽出により、有意差の認められた特徴について、それが認知症傾向の有無の判別に有用であるかを検証する。その方法と実験結果をまとめる。

判別方法アライメント情報をともなった対話音声およびその書き起こし文から得た特徴量のうち、検定により有意差が認められた特徴量を用いて識別を行うことにより、認知症傾向の有無を判別するのに有用であるかを検証する。識別に用いるデータは、データ数を音響的特徴に合わせて比較するため、アライメント情報を持つ22名(認知症傾向あり：9名、認知症傾向なし：13名)のものとする。識別には下記3種の特徴量を用い、その精度を比較・検討することで評価を行う。

- ・ openSMILE の解析結果による音響的特徴
- ・ 3章にて有意差が認められた言語的特徴(7次元)
- ・ 3章にて有意差が認められた音響的特徴(32次元)
- ・ 3章にて有意差が認められた言語的・音響的特徴(39次元)

表5 特徴量の種別による識別結果

	適合率	再現率	F値
言語的特徴 (7次元)	86.9%	86.4%	86.4%
音響的特徴 (32次元)	83.6%	77.3%	74.7%
言語的音響的特徴 (39次元)	88.9%	86.4%	85.7%

openSMILE とは、ミュンヘン工科大学が開発する「音声認識」「音楽認識」「パラ言語認識」の研究向けに作られたツールである。実行時に指定する構成ファイルにより、高次元の特徴量を抽出することができる[14]。openSMILEは韻律情報を用いた感情抽出や状態推定などによく用いられているため、今回は本研究で選択した特徴量の有用性を示すための比較対象として、openSMILE による解析結果を識別に用いた。openSMILE で利用可能な構成ファイルは多数存在するが、今回は IS09 という構成ファイルを比較対象として用いた。ただし、IS09 によって抽出される特徴量は 384 次元と、識別対象のデータ数に対して膨大であることから、主成分分析によって次元圧縮したものをを用いる。openSMILE による特徴を主成分分析にかけた結果、80%以上の累積寄与率をもつ主成分を採用した。第22 主成分までで累積寄与率が100%に達した。今回は、このデータを用いた識別結果を比較対象とした。識別にはデータマイニングツール Weka [15] を利用し、SVM(RBFkernel) で識別を行い、評価は10-foldcross validation で行った。なお、SVM(RBFkernel) のパラメータC, γ はパラメータ探索によって得た値を用いた。識別結果を表5 に示す。

言語的特徴あるいは音響的特徴を用いた結果よりも、言語的特徴および音響的特徴の両方を用いた結果が一番優れており、音響・言語両面から認知症傾向を判断することが有用であることが示された。しかし、認知症傾向を有している被験者のみから算出した再現率を確認してみると、言語的音響的特徴では66.7%と良い結果であったとは言い切れない。音響的特徴を用いた識別では、再現率が44.4%と非常に低い値であった。一方、言語的特徴では88.9%と高い再現率を示していることから、音響的特徴を用いた再現率の向上により、さらなる高精度化が期待される。

5. まとめ

本研究では、高齢者の対話音声を収集し、その言語的・音響的特徴を分析することで認知症傾向の判別を行うことを目的とした。大学、高齢者施設においてデータ収集を行い、男性22 名、女性76 名の計98 名からHDS-R 実施時の対話音声を収録した。このデータから計27 種の特徴を抽出し、認知症傾向別に検定にかけることで分析を行った。言語的特徴では、98 名(認知症傾向あり11 名、認知症傾向なし87 名)の対話音声の書き起こし文を用いて計18 種の特徴を抽出した。その結果、認知症傾向ありの被験者には代名詞割合の増加低下が認められ、失名詞失語による代名詞の増加が認められた。また、異なり名詞割合と語彙レベルには低下が認められ、認知症傾向ありの被験者は対話中に用いる名詞の種類が少なく、用いる語彙の難度も低くなる可能性が示された。対話中に含まれる品詞の割合では、特に認知症傾向ありの被験者の副詞割合が有意に低く、認知症傾向のある被験者は、日常的に交わされる対話中においても動詞や形容詞などの程度や状態を適切に示す表現力が乏しくなる可能性がある。また、感動詞割合においても低下が見られ、認知症の症状の一つである感情の発露やコミュニケーション意欲の低下が対話音声にも現れる可能性を指摘した。

音響的特徴では、22 名(認知症傾向あり9 名、認知症傾向なし13 名)の対話音声と音素アライメント情報を用いて計9 種の特徴を抽出した。その結果、音素別に32 種の特徴に有意差が認められ、特に子音のフォルマントからは認知症傾向のある被験者は調音の際の口腔内の狭めを作る動きに困難が生じる可能性を指摘した。

対話音声の言語的特徴・音響的特徴の内、有意差が認められた特徴を用いて識別を行い、認知症

傾向の判別に有用であるかの検証を行った。識別に用いたデータは、1)openSMILEで抽出した音響的特徴、2)有意差が認められた言語的特徴、3)有意差が認められた音響的特徴、4)有意差が認められた言語的・音響的特徴の4種で、openSMILEで抽出した音響的特徴は、比較対象として利用した。その結果、有意差が認められた音響的特徴を用いての識別、有意差が認められた言語的特徴を用いての識別は、ともに高い精度で認知症傾向を判別できていることが確認され、本研究で抽出した両特徴が認知症傾向の判別に有用であることが示された。さらに、言語的特徴と音響的特徴の両方を識別に用いた場合にはさらに精度が向上し、認知症傾向の判別において言語的・音響的両側面から特徴を抽出することが有用であることが示唆された。しかし、誤って識別された被験者は全て認知症傾向ありの被験者であるため、今後はこのような誤識別を減らすことが課題となる。

謝辞

本研究を遂行するにあたり、愛知県立大学 情報科学部の川島愛美氏および黒川有紀氏には多大な協力をいただきました。また、徳島大学大学院社会産業理工学研究部の北岡教英教授には高齢者音声のデータ収集および本研究の解析手法に関して有益なご助言をいただきました。ここに記して感謝申し上げます。

参考文献

- [1] 二宮利治, 清原裕, 小原知之, 米本孝二, 平成26年度厚生労働科学研究費補助金厚生労働科学特別研究事業「日本における認知症の高齢者人口の将来推計に関する研究」平成26年度 総括・分担研究報告書(2015).
- [2] 厚生労働省: 認知症施策推進総合戦略(新オレンジプラン)～認知症高齢者等にやさしい地域づくりに向けて～の概要, 認知症施策推進総合戦略(新オレンジプラン), 厚生労働省, 入手先<<http://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000064084.html>, (参照2017-10-10).
- [3] 大武美保子, "認知症の予防と支援に役立つ人工知能と高齢者とともにつくる認知症予防支援サービスの開発", 人工知能学会, 31(3), pp.363-370, 2016
- [4] 田中宏季, 足立浩祥, 浮田宗伯, 池田学, 数井裕光, 工藤喬, 中村哲: "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No.14, pp.1-4(2016).
- [5] 加藤昇平, 鈴木祐太, 小林朗子, 小島敏昭, 伊藤英則, 本間昭: "高齢者音声韻律特徴を用いたHDS-Rスコアとの相関分析", 人工知能学会論文誌 26 巻 2 号, pp.347-352(2010).
- [6] 国立情報研究所: 新聞記事読み上げ高齢者音声コーパス(S-JNAS), 音声資源コンソーシアム, 入手先<<http://research.nii.ac.jp/src/S-JNAS.html>>(参照2017-1-30)
- [7] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto: "Applying Conditional Random Fields to Japanese Morphological Analysis", Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004), pp.230-237(2004).
- [8] 松田実: "アルツハイマー型認知症の言語症状の多様性", 高次脳機能研究 第35 巻 第3号, pp.50-62(2015).
- [9] 大沢愛子, 前島伸一郎, 種村純, 関口恵利, 板倉徹: ""もの忘れ外来"における認知症と言語流暢性課題", 高次脳機能研究Vol.26 No.3, pp327-333(2006).

- [10] 玉井顯："認知症における高次脳機能検査について", 精神経誌,111 卷1 号,pp.93-100(2009).
- [11] Yuriko SUNAKAWA,Jae-ho LEE,Mari TAKAHARA："The Construction of a Database to Support the Compilation of Japanese Learners' Dictionaries",Acta Linguistica Asiatica,Vol. 2,No.2,pp.97-115(2012).
- [12] 原田大輔, 水町光徳, 二矢田勝行："高齢者のめりはりのない声に関する音響的解析", 電子情報通信学会,EA201e-86,pp.13-18(2010).
- [13] 石原健司："前頭側頭葉変性症にみられる発話障害の症候と病理", 高次脳機能研究 第31 巻 第3 号,pp277-283(2011).
- [14] Florian Eyben, Felix Weninger, Florian Gross, Björn Schuller: "Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor", In Proc. ACM Multimedia (MM), Barcelona, Spain, ACM, ISBN 978-1-4503-2404-5, pp. 835-838, October 2013. doi:10.1145/2502081.2502224
- [15] S. R. Garner："WEKA: the Waikato environment for knowledge analysis", Proceedings of the New Zealand Computer Science Research Students Conference, pp.57-64(1995).