音声発話による認知症スクリーニング技術の開発
感情表現辞書を用いた発話内容の質的分析

柴田 大作[1)]　若宮 翔子 Ph.D.[1)]　木下 彩栄 Ph.D. MD[2)]　荒牧 英治 Ph.D.[1)]

1) 奈良先端科学技術大学院大学 情報科学研究科　2) 京都大学大学院 医学系研究科

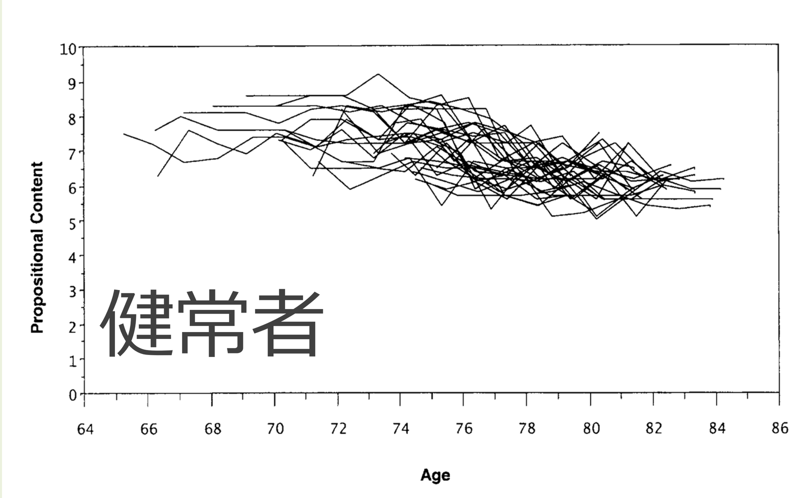
1. 背景

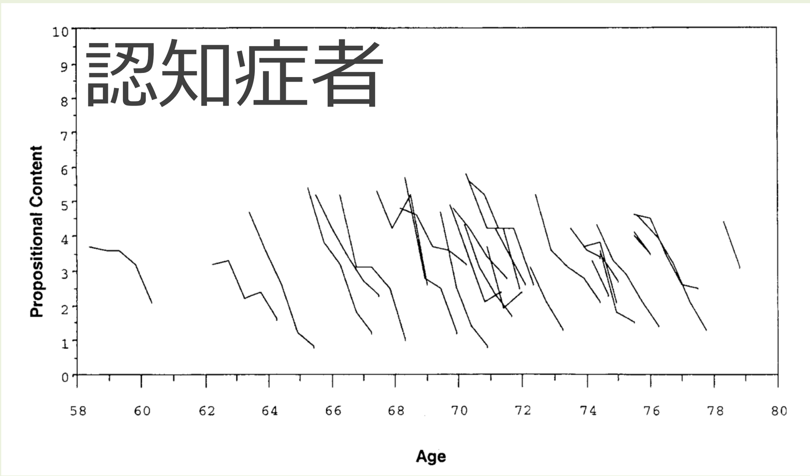
高齢者の運転免許更新における**認知機能検査**の義務化

⇩

認知症の**早期発見**は重要な課題

認知症患者の**語彙能力**は急激に低下





[Snowdon+ 1996]

語彙能力に着目した認知症のスクリーニングを検討

2. 従来手法の問題点

従来手法

- 語彙量 [Fraser+ 2016] [Calvin+ 2016]
- 品詞数 [Jarrold+ 2014] [Fraser+ 2016]
- フィラー数 [Tanaka+ 2016]
- 冗長性 [Aramaki+ 2016]　など

基本的には患者の状態を1つの統計量にまとめ込む

⇩

患者の多様な発話を1つの統計量として扱うことは
問題を過剰に簡単化している恐れ

感情表現辞書を用いた
質的な観点を残しながら統計量により測定

3. 材料

研究参加者のリクルート：京都大学医学研究科

認知症者/健常者の語りの材料として，発話を録音し書き起こしたコーパスを材料として利用

1. 心理検査コーパス（京都大学「医の倫理委員会」承認番号：E2525）

受診中の患者が心理検査時に話す内容を書き起こしたコーパス

AD: アルツハイマー型認知症 HC: 健常高齢者

【包括基準】

- 認知症群：MCIを含む軽度から中程度のアルツハイマー型認知症患者（MMSE21点以下）
- 健常者群：認知症でないと確認でき中枢神経系に異常を認めない認知症群と年齢をマッチさせた群（MMSE22点以上）

【除外基準】

- 中枢神経系の代謝に影響を及ぼす疾患を持つ者
- 病状などにより十分な同意能力を持たない者
- 非日本語母語者

	AD	HC
	認知症と診断され，MMSEの結果が21点以下	認知症でないと診断され，MMSE結果が22点以上
性別	男性: 2人 女性: 7人	男性: 4人 女性: 5人
年齢（平均）	77.2	76.6
MMSE（平均）	17.0	25.1

いや、みて見てなかったです。……えっと…全く見てなかったですね。季節はどっちか言うたら、春でしょう。
京都の病院。いや、分かりません。４階くらいじゃなかったですか？こ、こ、京都市になりますかね？
京都府になります。京都のほうだから、京都は言い方あったかな？……京都は言い方無かったと思いますけどね。桜、猫、電車。

2. 自由会話コーパス（京都大学「医の倫理委員会」承認番号：C1152）

以下の質問における回答内容を書き起こしたコーパス

1. 昔楽しかったことはなんですか
2. 最近楽しかったことはなんですか
3. 好きな食べ物はなんですか

Dementia: 認知症 Mild Cognitive Impairment (MCI): 軽度認知障害

【包括基準】

- 認知症の診断を受けている65歳以上の高齢者
- MMSE10点以上の軽度～中程度の認知症患者
- 「さびしいと感じることがありますか」などの質問に「はい」「いいえ」で答えることができる者
- 関西方言話者

【除外基準】

- 認知症の診断がない者
- MMSE10点未満の重度認知症患者
- 意思疎通に障害を有する者
- 重度な聴覚障碍者
- 関西方言以外の方言話者
- 認知症に対する治療が新たに開始、または変更になる予定がある者

	Dementia	MCI
	認知症と診断され，MMSEの結果が21点以下	認知症と診断され，MMSE結果が22点以上
性別	男性: 4 女性: 19	男性: 3 女性: 16
年齢（平均）	86.17	82.37
MMSE（平均）	15.00	25.58

<イ>
はいはい
<高>
はぁ。もうきちんとしてくれますしね、それぐらいですね。
家でだったらね、独りで自分でしないといけないか、とって食べないといけないからね
とって食べたら味なかったしね。前も…
ここだったらね、きちんとしてくれますしね。
はぁ。……
<イ>
分かりました、では、好きな食べ物は何ですか？
<高>
私、好きな食べ物。……なんでも食べるけどね。……。
嫌いなものって無いけどね、私。ふん。
あの、<笑「ふ」>鮒ずしとか嫌いよ、
滋賀県のね、臭いしね、主人は好きだったね、滋賀県だからね。……

4. EPISODE BANKと日本語感情表現辞書（JIWC）の構築

EPIOSODE BANKの構築

EPISODE BANK

クラウドソーシングにより，いくつかの感情に関するエピソード記憶を収集し、EPISODE BANKと呼ばれるデータセットを構築

収集する感情はプルチックの感情の輪 [Plutchik 2001]に基づき設定し、タスク内容（質問形式）はイギリスのコホート研究[Golding+ 2014]を参考に設定

1000人のクラウドワーカーから7000個のエピソード記憶を収集
データーは本研究室のホームページで公開を予定

No	タスク内容
1	以下の単語から思い出されるあなたの人生における最も印象的な出来事を教えてください。「悲しい」
2	以下の単語から思い出されるあなたの人生における最も印象的な出来事を教えてください。「不安」
3	以下の単語から思い出されるあなたの人生における最も印象的な出来事を教えてください。「怒り」
4	以下の単語から思い出されるあなたの人生における最も印象的な出来事を教えてください。「嫌悪感」
5	以下の単語から思い出されるあなたの人生における最も印象的な出来事を教えてください。「信頼感」
6	以下の単語から思い出されるあなたの人生における最も印象的な出来事を教えてください。「驚き」
7	以下の単語から思い出されるあなたの人生における最も印象的な出来事を教えてください。「楽しい」

回答例

No.1: 悲しい
- 両親が立て続けに亡くなったこと
- かわいがっていた愛犬が死んだ時
- 好きな女の子に告白してふられたとき

No.2: 不安
- 大地震を体験したこと
- 就職の面接の時
- 20代で阪神淡路大震災に遭遇した

No.3: 怒り
- 浮気をされたこと
- 小学生のとき初デートのタイミングで雨が降って中止になっとき，天候にイカリ
- 友人と思っていた人間に裏切られたとき

No.4: 嫌悪感
- 浮気をされたこと
- 大嫌いな上司に嫌味を言われたとき
- うそをつかれていた時

No.5: 信頼感
- いい人と結婚できたこと 困ってる自分を俺が助けてやる、と言ってくれた
- 小学生の時の担任の先生
- 両親が困ったときにいつでも味方になり手助けをしてくれる

No.6: 驚き
- 子供が受験に合格したこと
- 大学に合格した時
- 自分と同性同名の人を調べたら、全国にすごく多くいたこと。

No.7: 楽しい
- 自立できたこと
- バスケットボールの試合をしているとき
- 久しぶりに遠くに住む家族に逢えて嬉しく、楽しい日を過ごせた時があった。今は少し疎遠がちに。

日本語感情表現辞書（JIWC）の構築

;): JIWC
Japanese linguistic Inquiry and Word Count

EPISODE BANKのエピソード記憶のデータセット（以下、データ）を利用

Step1. データの形態素解析

データの形態素解析を（Juman++）により実行
各形態素の代表表記を取得

Step2. 単語頻度・割合・スコアの算出

全回答データの形態素と設問ごとの回答データの形態素を比較し，各感情表現の単語頻度を算出
こ各単語の合計頻度が10以下の単語は削除
各感情における単語頻度と単語割合を算出

Step3. 感情表現の抽出

Step2の単語頻度，単語割合を用いて単語スコアを算出

$$\mathrm{Score} = \frac{w_{ij}}{w_i *}\log(w_{ij}+1)$$

Score: 単語スコア,
w_{ij}: 単語w_iの感情jでの頻度
w_{i*}: 単語w_iの全ての感情での頻度

Step4. 感情表現の抽出

単語ごとに単語スコアが最も大きくなる感情を求め，その感情に対応する表現として単語を辞書に追加

単語スコアが大きい上位20単語

悲しい	不安	怒り	嫌悪感	信頼感	驚き	楽しい
亡くなる	不安だ	怒り	嫌悪	くれる	驚く	楽しい
死	将来	理不尽だ	痴漢	信頼	驚き	旅行
悲しい	受験	怒る	セクハラ	助ける	宝籤	遊ぶ
死ぬ	老後	喧嘩	不潔だ	困る	当たる	趣味
祖母	就職	れる	抱く	アドバイス	懸賞	ライブ
別れ	失業	裏切り	汚い	パートナー	当選	出掛ける
祖父	転職	騙す	臭い	相談	妊娠	行く
飼う	今後	カード	感	感	成長	新婚
ペット	結果	裏切る	嫌だ	任す	合格	過ごす
死去	面接	他人事	隣り	任せる	みる	ディズニーランド
他界	入試	られる	電車	支える	プレゼント	コンサート
可愛がる	試験	元	奴	辛い	意外だ	イベント
失恋	活動	者	しつこい	先生	結婚	アーティスト
愛犬	生活	勝手だ	迷惑	心	高い	美味しい
犬	入院	逃げる	座る	心配	息子	家族
婆ちゃん	発表	暴力	態度	出会い	出産	デート
飼い犬	決まる	嘘	嫌いだ	何でも	告白	休み
祖父母	見付かる	詐欺	嫌味だ	もらう	急だ	遊び
死別	病気	突ける	人	関係	テレビ	時間
亡くす	貯金	扱い	行為	妻	子供	飲む
父	進路	馬鹿だ	男	守る	知り合い	野球
葬式	リストラ	追う	人間	悩み	思う	好きだ
しまう	手術	苛め	言動	持つ	東京	夏

5. JIWCによる認知症のスクリーニング

手順

Step1. コーパスの形態素解析

形態素解析機器：(Juman++)
コーパス中の代表表記を取得

Step2. JIWCによる分析

代表表記とJIWCをマッチ
各感情の使用頻度・割合を算出

例文:「大学受験に合格したこと」
⇩ **Step1**: 形態素解析により代表表記を取得
大学/受験/合格
⇩ **Step2**: JIWCとのマッチング
大学　受験　合格
<不安> <不安> <驚き>

t検定による分析

（A）心理検査コーパスの分析結果

	AD群平均値	HC群平均値	*p-value*
悲しい	0.037	0.054	0.052
不安	0.087	0.059	**0.049***
怒り	0.141	0.101	**0.005***
嫌悪感	0.080	0.060	**0.009***
信頼感	0.135	0.110	0.109
驚き	0.074	0.059	0.173
楽しい	0.060	0.071	0.320

感情「不安」・「怒り」・「嫌悪感」の使用割合がAD群で有意に増加

（B）自由会話コーパスの分析結果

	Dementia群平均値	MCI群平均値	*p-value*
悲しい	0.070	0.063	0.508
不安	0.081	0.052	**0.004***
怒り	0.129	0.125	0.766
嫌悪感	0.103	0.093	0.404
信頼感	0.128	0.147	0.182
驚き	0.058	0.063	0.582
楽しい	0.130	0.146	0.318

感情「不安」の使用割合がDementia群で有意に増加

6. 考察

（A）心理検査コーパスの考察

- アルツハイマー型認知症の周辺症状として，**怒りや不安**を感じるということが，一般的に言われている特徴と一致
- 日本語感情表現辞書により，**定性的な特徴を定量的に評価**することができる可能性
- 感情「嫌悪感」については現時点で十分な考察ができておらず，今後の課題としたい

（B）自由会話コーパスの考察

- 協力施設内で認知症高齢者の方に喋ってもらうというスタイルで自由会話データの収録を行ったため，施設で普段行わないイベントが発生し，**不安を感じた可能性あり**
- 参加者の「どうして知らない人間に話さなければならないのか」などといった漠然とした感情が**語りに現れた可能性**

7. まとめ

- クラウドソーシングにより日本語感情表現辞書（JIWC）を開発
- JIWCによる発話の質的部分に着目した認知症スクリーニングの可能性について調査
- Dementia群とAD群の両群において感情「不安」を表現する単語の使用割合が有意に増加

8. 今後の課題

- 他の精神的疾患（発達障がいやうつ病）に適用し，スクリーニングの可能性を検討
- 日本語感情表現辞書の更新

謝辞

研究参加者のリクルート，データの収集などにおいて終始ご協力いただきました京都大学大学院医学研究科の鳥畑由香里さまに厚く御礼申し上げます。