

アルツハイマーの発症に伴う代名詞の増加

柴田大作^{†1} 若宮翔子^{†1} 木下彩栄^{†2} 荒牧英治^{†1}

概要：近年、認知症の早期発見が社会的にも大きな意味を持ち始めている。そこで、症状の一つである認知能力の低下に伴う言語能力の低下を調査する研究が注目されている。本研究は、認知症者の発話における単語の特徴を精査することで、認知症者のスクリーニングを実現する。本稿では、MMSE テストによりアルツハイマー型認知症患者（8名）と健常高齢者（8名）の対話を書き起こし、単語をカテゴリに分類した辞書である LIWC を用いて、発話におけるカテゴリ出現頻度を調査する。認知症群において、代名詞カテゴリの出現頻度が有意に多いことが確認された。従来から認知症患者は「これ」や「あれ」等の代名詞を多用することが報告されており、本研究においてその事実を定量的に示すことができた。今後、発話中の単語のカテゴリ出現頻度に基づく効率的な認知症スクリーニング技術の確立が期待される。

Pronoun Increases in Alzheimer's Disease

DAISAKU SHIBATA^{†1} SHOKO WAKAMIYA^{†1}
AYAE KINOSHITA^{†2} EIJI ARAMAKI^{†1}

1. はじめに

世界に先駆けて高齢化社会を迎える日本にとって、高齢者への医療対策は重要な課題である。中でも認知症は、社会的費用が年間 14.5 兆円にのぼるとも算定されており、国家的な問題となっている。現在、認知症の予防または進行を抑制する手法の確立に向け、大規模な臨床試験が世界各地で開始されつつある。近い将来、認知症を予防できる可能性が示されれば、次は、いかにして認知症を早期発見するかが課題になると想定される。

早期に症状を発見する研究については、近年、血液検査など新たな方法が提案されている。しかし、それらは身体的侵襲を伴うものが多い。そこで、この負担を低減するため、認知症の主な症状である記憶能力や言語能力の低下に焦点をあてた調査・研究が注目されている。

今日に至るまで認知症のスクリーニングに関する研究は数多くなされてきた。言語に注目した研究は 2007 年頃 [1] から始められており、最近では発話サイズに注目したものが報告されている [2]。そして、言語のみならず音声に焦点をあてた研究 ([3], [4], [5], [6]) や、神経テスト [7] を用いた研究もある。これらを表 1 にまとめる。

このように認知症スクリーニングに関する研究は多く報告されているが、基本的に患者の状態を一つの統計量にまとめこむものが多い。患者の多用な発話を一つの統計量として扱うのは簡単ではあるが、過剰に問題を単純化してしまう恐れがある。例えば、認知症者についてよく言われているような「不平をよく言う」などの特徴を、先行研究で用いられている語彙量や冗長性により観察することは難しい。

本研究では、「不平」のような質的な概念を測定するために、語彙カテゴリごとの集計を用いた認知症のスクリーニング法を提案する。そのために、単語がカテゴリ毎に分類された辞書である LIWC (Linguistic Inquiry and Word Count) を利用する。認知症者と対照者が発話中に使用した語をカテゴリ毎に集計し、両者におけるカテゴリの使用割合に差があるかを調査する。これまでも認知症者 (AD) とそうでない者 (HC) が使用する言語の違いに関しては報告されてきた。例えば、一般的に認知症者は「それ」や「あれ」などの非人称代名詞を多用することが報告されている。これらは定性的には検証されてきたが、定量的には検証されてこなかった。そこで、これらに関して調査し、また他のカテゴリに関しても両者の間に差があるのかを比較・検討する。

2. 材料

認知症者の語りを材料として、病院を受診中の患者が心理検査時に話す内容を収録し、それを書き起こしたテキストを用いる。

^{†1} 奈良先端科学技術大学院大学
〒630-0192 奈良県生駒市高山町 8916-5 奈良先端科学技術大学院大学 研究推進機構 ソーシャル・コンピューティング研究室

^{†2} 京都大学医学研究科 人間健康科学系専攻
〒606-8507 京都府京都市左京区聖護院川原町 53

表 1. 認知症のスクリーニングに関する先行研究.

	概要	疾患	サンプルサイズ	年
Aramaki et al. [2]	作文とスピーチの語彙量の解析	MCI, AD	22	2016
Tanaka et al. [3]	コンピュータアバターを用いた音声対話の解析	MCI	18	2016
König et al. [4]	スピーチの解析	MCI, AD	64	2015
Tóth et al. [5]	スピーチの解析	MCI	51	2015
Moriyama et al. [7]	神経テスト	AD	299	2015
Roark et al. [6]	スピーチ (書き起こし) の解析	MCI	74	2011
Roark et al. [1]	スピーチ (書き起こし) の解析	MCI	55	2007

2.1 フィールド

研究協力者のリクルートは、京都大学医学部附属病院神経内科にて行った。これには以下の基準を用いた。

【包含基準】

- 認知症群：症状の前兆の軽度認知障害 (MCI), レベル軽度から中等度までのアルツハイマー型認知症患者 (MMSE21 点以下)
- 対照者群：認知症でないこと確認でき、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ような疾患を持つ者、病状などにより十分な同意能力を持たない者。非日本語母語話者。

2.2 倫理・リスクに対する配慮など

研究協力者、および、代諾者 (保護者) 同席の場合には代諾者に対して別紙説明文に記載された内容について十分に説明を行った。病状のため、研究協力者と代諾者が共に研究内容についての説明を十分に理解することができないと判断される場合は、研究対象から除外した。なお、患者へは報酬などの利益は支払っていない。本研究は、京都大学「医の倫理委員会」の倫理審査の承認を経て実施された。

3. 日本語版 LIWC の作成

LIWC は語彙をカテゴリ化する際に使用されるツールである。これは社会学、認知心理学や臨床などの分野の研究者によって開発されたものであり、人々の社会的な状態や心理的な状態を、語彙によりカテゴリ化することが可能である。しかし、LIWC は全て英語であり、日本語で作成された文章に対して用いることは難しい。そのため、本研究ではまず LIWC の日本語化を行う。

表 2. LIWC2007 の全カテゴリ.

<Funct>	<Sad>	<Adverbs>	<Body>
<Pronoun>	<CogMech>	<Prep>	<Health>
<Ppron>	<Insight>	<Conj>	<Sexual>
<I>	<Cause>	<Negate>	<Ingest>
<We>	<Discrep>	<Quant>	<Relativ>
<You>	<Tentat>	<Numbers>	<Motion>
<SheHe>	<Certain>	<Swear>	<Space>
<They>	<Inhib>	<Social>	<Time>
<Ipron>	<Incl>	<Family>	<Work>
<Article>	<Excl>	<Friends>	<Achiev>
<Verbs>	<Percept>	<Humans>	<Leisure>
<AuxVb>	<See>	<Affect>	<Home>
<Past>	<Hear>	<Posemo>	<Money>
<Present>	<Feel>	<Negemo>	<Relig>
<Future>	<Bio>	<Anx>	<Death>
<Filler>	<Nonflu>	<Anger>	<Assent>

表 3. 疾患と関係があると判定されたカテゴリ.

<Time>	<We>	<Past>	<I>
<Future>	<Family>	<Ipron>	<They>
<Social>	<SheHe>	<Anger>	<Present>
<Posemo>	<Anx>	<Verbs>	<Friends>
<Space>	<Negemo>	<Sad>	<You>
<Negate>	<Humans>		

3.1 前処理

日本語翻訳する前処理として、カテゴリの整理を行う。表 2 に LIWC2007 の全カテゴリを示す。ここで、疾患と関係がないと思われるカテゴリ (例: <Body>) や日本語では扱われないカテゴリ (例: <Article>) を筆者らの判断により削除した。これにより、カテゴリ数は 64 から 22 となった。これを表 3 に示す。

3.2 LIWC の翻訳

以下の手順に沿って LIWC の日本語版を作成した。

Step1 : 日本語翻訳

電子辞書 (EDP 製: 英辞郎, Jim Breen: edict) を用いて, 前処理により残ったカテゴリ (表 3) の単語を自動的に日本語訳した。

Step2 : 誤訳によるノイズ除去

明らかに不適切な日本語訳がなされた単語はノイズである。作業員 1 名が目視によりデータを確認し, ノイズを削除した。これにより総単語数が 6,211 語から 5,534 語となった。誤訳の例を表 4 に示す。

Step3 : 重複除去

異なる英単語でも同一の日本語に翻訳される場合がある。そこで, カテゴリと日本語が一致するものは一つにまとめた。この例を図 1 に示す。これにより, 総単語数が 5,534 語から 4,769 語となった。

Step4 : カテゴリの再割当

日本語の動詞には過去形が存在しないことから, <Past> のカテゴリに属する動詞を全て削除した。また, <Present> に属する動詞のカテゴリを <Verbs> に変更した。そして, カテゴリとの関係を人手で判断して, 関連性が低いと思われる 3 つのカテゴリ (<We>, <SheHe>, <They>) を削除した。これによりカテゴリ数が 22 から 19 になった。

Step5 : 新しいカテゴリの定義

複数のカテゴリに属する語に関しては, 作業員 1 名が目視により正しいカテゴリ (最大一つ) を選択した。また, 日本語では時間と空間の定義は曖昧である (表 5)。そのため, <Time> (時間) と <Space> (空間) を一つにまとめ, 新たなカテゴリとして <TimeSpace> (時間・空間) を定義した。これによりカテゴリ数が 19 から 20 になった。

LIWC ファイルのうち, google n-gram コーパス上で出現頻度の多かった 2,700 語に対し, Step1 から Step5 までの作業を適用した。日本語版 LIWC における最終的なカテゴリを表 6 に示す。各カテゴリに属する英単語と日本語の例を表 7 に示す。

表 4. 誤訳の例.

カテゴリ	英単語	日本語
Space	Night	すぐ
Posemo	Love	じ

表 5. カテゴリの曖昧な語.

カテゴリ	英単語	日本語
Time	While	中
Space	in	中

カテゴリ	英単語	日本語
<Humans>	girl	娘
<Humans>	chick	娘



カテゴリ	英単語	日本語
<Humans>	girl, chick	娘

図 1. LIWC の翻訳におけるカテゴリのまとめあげ.

表 6. 日本語版 LIWC のカテゴリ.

<Time>	<Verbs>	<Space>	<Family>
<Future>	<Humans>	<Past>	<TimeSpace>
<Social>	<Sad>	<Friends>	<Anx>
<Present>	<Negate>	<Ipron>	<Negemo>
<Posemo>	<You>	<Anger>	<I>

表 7. 日本語版 LIWC のカテゴリ.

カテゴリ	英単語	日本語
<Time>	end, next, age	終了, 次, 時期
<Future>	will, might, wish	決意, 勢力, 願う
<Social>	mail, help, advice	メール, ヘルプ, 勧め
<Present>	support, use, make	支援, 使用, 作成
<Posemo>	good, hope, lovely	良い, 希望, 可愛い
<Space>	town, in, land	町, 中, 地域
<Past>	origin, old, used	元, 古い, 中古
<Friends>	mate, companion, boy	友達, 連れ, 若者
<Ipron>	stuff, this, who	物, 当, 誰
<Anger>	war, cut, temper	戦, 削除, 気分
<Verbs>	be, hit, doing	いる, 打つ, 行い
<Humans>	girl, male, woman	娘, 男性, 女性
<Sad>	suffer, sadness, cry	悩む, 悲しみ, 泣く
<Negate>	without, not, nope	なし, ない, いや
<You>	you	君, あなた, おまえ
<Family>	family, child, folks	家族, 子, 皆さん
<TimeSpace>	before, right, while	手前, 直ぐ, 中
<Anx>	fear, terrifying, turmoil	心配, 怖い, 混乱
<Negemo>	dislike, bad, awful	嫌い, 悪, ひどい
<I>	I	おれ, ぼく, わたし

3.3 配布場所

本研究により作成した日本語版 LIWC は, 国立大学法人 奈良先端科学技術大学院大学ソーシャル・コンピューティング研究室のホームページ^aより入手可能である。

4. 実験

4.1 言語リソース

本研究では, 認知症者の語り (n = 18, AD = 9, HC = 9) をコーパスとして使用する。図 2 に一部抜粋したものを示す。これは 1 名の発話テキストの約 14% の文量である。

a <http://sociocom.jp/>

あんまり感じないんですけども、その、あの、
 ひそひそ声ではしる、は、早、あの話す人が家には
 無いので、
 主人もまーどっちか言ったら、
 ちょっと遠いくらいなってきたし
 (中略)
 はい。
 はい、主人です。
 なんかあの、息子も行ってみるってゆうてはいたら
 しいんですけど、
 電話で、まだ来てなかったですし、
 おい、お医者してますので、

図 2. 発話の例.

4.2 手順

以下の手順に沿ってコーパスを解析した.

Step1: コーパスに対して形態素解析を行い、全単語の形態素を取得する.

Step2: Step1 で取得した全単語の形態素と日本語版 LIWC の全単語を比較し、コーパス中の単語の使用数をカウントする.

Step3: 得られた単語の使用数をカテゴリの使用数に変換することで、会話における各カテゴリの使用割合を取得する.

Step4: Step1 から Step3 の作業を、全コーパス (n=18) に対して適用した.

4.3 結果

認知症群 (AD) と対照者群 (HC) の発話から得られた各カテゴリの使用割合のばらつきを図 3 に示す. また、AD と HC の各カテゴリの平均使用割合の差が統計的に有意であるかを検定するために、有意水準 $p < 0.05$ で t 検定 (片側検定, 等分散を仮定しない) を行った. この結果を表 8 に示す.

この結果より、<Time> (時間)、<Space> (空間)、<Past> (過去)、<Friends> (友人)、<Anger> (怒り)、<Humans> (人)、<Sad> (悲しみ)、<Negate> (否定)、<You> (二人称)、<Family> (家族)、<TimeSpace> (時間・空間)、<Negemo> (ネガティブな感情)、<Posemo> (ポジティブな感情) ならびに <I> (一人称) に関するカテゴリに関しては認知症群と対照者群の発話での使用頻度に統計的に有意な差は見られなかった. しかし、<Social> (社会)、<Present> (現在)、<Ipron> (非人称の代名詞)、<Anx> (心配・不安)、<Verbs> (動詞) に関しては、有意差があることが確認された. 表 9 に、有意差が確認された 5 つのカテゴリ <Social>、<Present>、<Ipron>、<Anx> および <Verbs> に属する単語の発話中での使用例を示す.

表 8. AD と HC の発話における各カテゴリの平均使用割合の差に対する t 検定の結果. 太字かつ下線の値は有意水準 $p < 0.05$ で統計的な有意差が確認されたもの.

カテゴリ	AD	HC	p
<Ipron>	0.039	0.027	0.019
<Anx>	0.001	0.000	0.019
<Verbs>	0.052	0.043	0.022
<Present>	0.017	0.010	0.023
<Social>	0.006	0.012	0.023
<I>	0.004	0.002	0.059
<Space>	0.017	0.023	0.089
<Posemo>	0.006	0.008	0.125
<Time>	0.036	0.042	0.143
<Sad>	0.000	0.000	0.173
<You>	0.000	0.000	0.269
<Family>	0.002	0.002	0.314
<Negate>	0.040	0.046	0.326
<Negemo>	0.001	0.001	0.329
<Anger>	0.000	0.001	0.339
<Humans>	0.007	0.008	0.343
<Friends>	0.001	0.001	0.402
<Non-Category>	0.771	0.775	0.422
<Past>	0.000	0.000	0.491
<Future>	0.000	0.000	-
<TimeSpace>	0.000	0.000	-

表 9. 有意差が見られたカテゴリの単語を含む発話例.

カテゴリ	使用例
<Ipron>	ここではまた違うけどね。 あの、言われたりするんですけど何かね、お互いね、
<Anx>	怖い怖い。 心配したら余計それがな、
<Verbs>	その後、ちょっと音楽大学に行ったんで。 英語はわかるんですけどよね。
<Present>	珍しい事あるさかい 今言いましたでしょ？
<Social>	ちゃんと聞いてかないかんから いうて言ってくれはったん。

5. 考察

本章では、まず結果について考察をする (5.1 節). 次に、決定木を用いて本研究の応用の可能性について議論する (5.2 節). 最後に、医学分野で一般的に用いられる方法に従い、ROC 曲線による考察を加える (5.3 節).

5.1 得られた知見: 先行研究の量的な裏付け

実験で有意差が見られたカテゴリに関して考察する. <Social> (社会) の使用割合は、AD で有意に低くなっている. 既存の認知症研究において、社会活動が少ない人は認知症を発症しやすくなるため、認知症の予防には社会活動が重要であることが報告されている [8]. さらに、認知症患者は社会活動が少ないという報告 [9] もある. このように、今回の実験結果は、一般的に言われている認知症の特徴と一致する.

一方、<Ipron> (非人称の代名詞) に関しては、AD の使用割合が有意に高くなっている. 認知症患者は物の名前を想起することが困難で、代名詞を多用することが既存

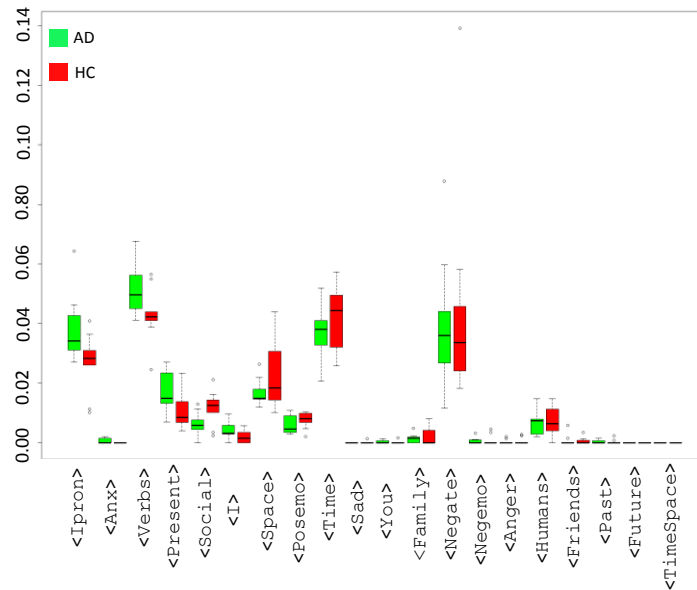


図 3. AD と HC の発話におけるカテゴリ使用割合のばらつき。<Ipron>, <Anx>, <Verbs>, <Present>, <Social> の 5 つのカテゴリに統計的有意差が確認された。

研究で報告されている [10]. このことから、<Social> の場合と同様に、この結果はこれまで指摘されている認知症の特徴と一致する。

同様に、<Anx> (心配・不安) の使用割合は、AD で有意に高くなっている。これは、認知症の症状 (記憶障害や被害妄想など) により、不安を感じるが多くなるのが原因ではないかと考えられる。なお、AD と鬱の初期症状は類似しており、関連が指摘される [11].

さらに、<Verbs> (動詞) と <Present> (現在) に関しても同様に、AD での使用割合が有意に高くなっている。これらに関しては今後、研究を進める過程で原因を明らかにしていきたい。

まとめると、発話における語のカテゴリ頻度を調査するアプローチにより、一般的に言われている認知症の症状と一致する結果が得られた。また、先行研究のような語彙量を調査するだけでは難しい、発話における質的な概念の測定による認知症患者のスクリーニングが可能であることが示された。

5.2 機械学習の適用: 決定木による考察

次に、単一のカテゴリだけでなく、カテゴリの組み合わせにより AD と HC の鑑別を試みる。そのために、各群の発話で用いられたカテゴリの割合を特徴量とした決定木を用いる。その結果、生成された決定木を図 4 に示す。図 4 より、<Ipron> (非人称の代名詞) の割合が 0.0298 以下かつ <Posemo> (ポジティブな感情) の割合が 0.0046 以下であれば、100% の確率で認知症であったことが分かった。これは、<Ipron> の使用頻度が少なくても <Posemo> の使用頻度を確認することで認知症患者のスクリーニングが

可能であることを示している。なお、<Posemo> は有意差こそ確認されなかったが、認知症患者群で少ない傾向にある。つまり、単一のカテゴリだけで判断するのではなく、複数のカテゴリを組み合わせることにより、スクリーニングの精度をより向上させることができる。

また、<Ipron> (非人称の代名詞) の割合が 0.0298 より大きく、<Space> (空間) の割合が 0.0025 以下、かつ <Posemo> (ポジティブな感情) の割合が 0.0025 より大きければ、同様に 100% の確率で認知症であったことが分かった。

これらの結果より、発話における <Ipron>, <Space>, および <Posemo> に関する単語の使用割合が、認知症のスクリーニングにおいて重要な指標であることが明らかとなった。

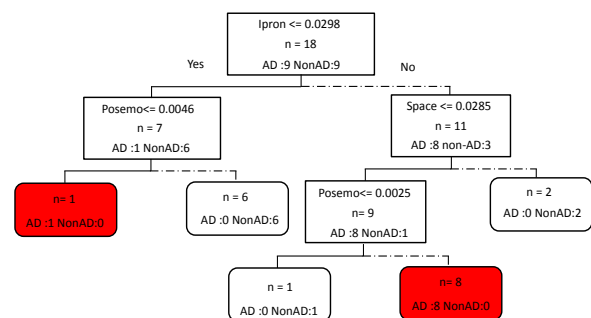


図 4. AD と HC のカテゴリ使用割合を特徴量として生成された決定木。発話における <Ipron>, <Space>, および <Posemo> に関する語の使用割合が認知症のスクリーニングにおいて重要であることが分かる。

5.3 医学的統計: ROC 曲線による考察

ROC 曲線による考察を行う。図 5 に、4.3 節の t 検定において有意差が示されたカテゴリの ROC 曲線を示す。ROC 曲線は、縦軸に感度、横軸に偽陽性率(=1-特異度)をとる。ROC 曲線下の面積 (Area Under Curve; AUC) により、分類アルゴリズムの性能を評価することができる (0 から 1 までの値を取り、ランダムな分類の場合は 0.5、完全な分類が可能な場合は 1 となる)。

<Ipron> (非人称の代名詞)、<Present> (現在) ならびに <Verbs> (動詞) に関しては、AUC 値が妥当な性能の下限値 (0.7) を超えており、信頼できると考えられる。しかし、<Anx> (心配・不安) に関しては、認知症患者の中でのみ使用されており AUC 値も妥当な性能下限を超えてはいるが、9 人中 4 人と母数が少ないため、信頼に欠ける。

一方、<Social> (社会) の AUC 値は 0.765 (1-0.235) と十分な性能をもっている。これは、認知症患者は <Social> の語の使用割合が少ないことを示している。すなわち、<Social> の使用割合と認知症に何らかの相関があることが示唆されている。

ROC 曲線の結果から、t 検定により有意差が確認されたカテゴリ (<Anx>を除く) に関しては、妥当な判断性能を持つことが示された。今後、<Social>、<Ipron>、<Present> および <Verbs> の語の使用頻度に着目した認知症のスクリーニングが期待される。

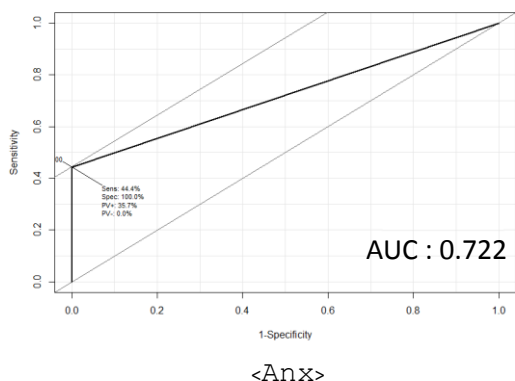
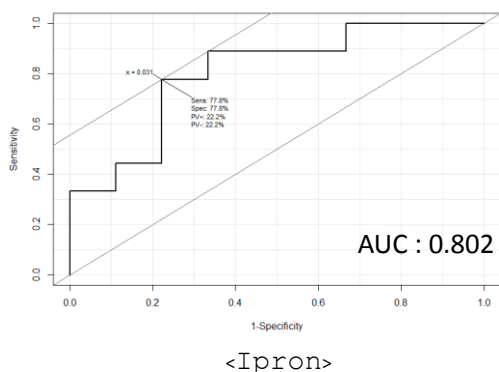
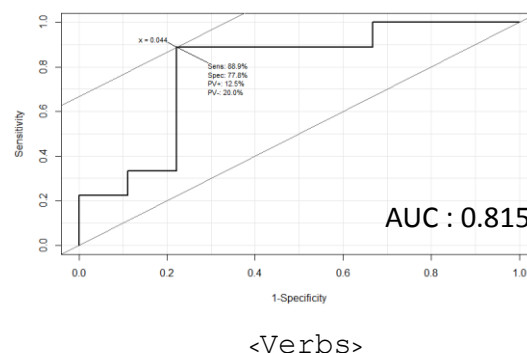
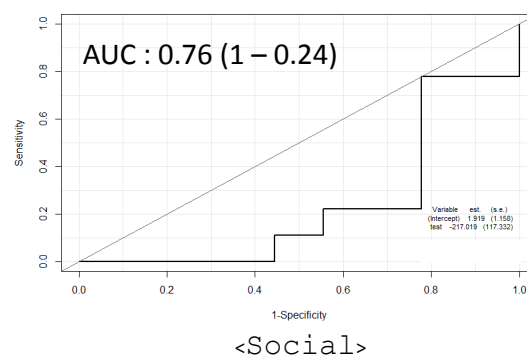
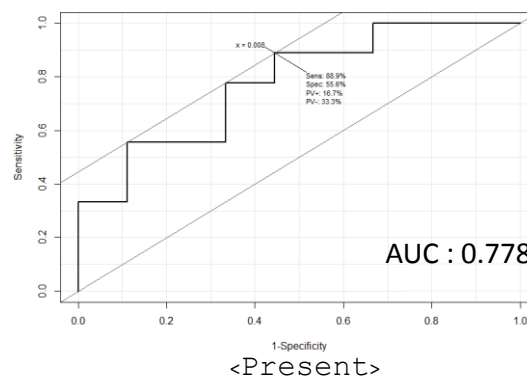


図 5. 有意差が見られたカテゴリの ROC 曲線。 <Ipron>、<Present> および <Verbs> の AUC 値は 0.7 を上回る。 <Social> の AUC 値は大幅に下回っており、妥当な判断性能を持つといえる。

6. おわりに

本研究は、自然言語処理技術を用いて、患者との対話をデータ化し、形態要素解析を行い、単語単位に分解した。そして、LIWC と比較し、カテゴリの一致率を調べ、認知症患者の語彙のカテゴリの偏りを調査した。結果、一部のカテゴリにおいて有意に患者を区別することができた。特に、<Ipron>は AUC 値が 0.802 と認知症を一定の精度でスクリーニングすることが可能であることを示した。今後は、音声入力ソフトウェアを用いることでさらに簡単にスクリーニングを実現するために発展させていく。

参考文献

1. Roark, B., M. Mitchell, and K. Hollingshead. Syntactic complexity measures for detecting mild cognitive impairment. in Proceedings of the Workshop on BioNLP 2007: Biological, Translational, and Clinical Language Processing. 2007. Association for Computational Linguistics.
2. Aramaki, E., et al., Vocabulary Size in Speech May Be an Early Indicator of Cognitive Impairment. PLoS One, 2016. 11(5): p. e0155195.
3. Tanaka, H., et al., <Automatic Detection of Very Early Stage of Dementia through Spoken Dialog with Computer Avatars.pdf>, in International Conference of the IEEE Engineering in Medicine and Biology Society. 2016.
4. Konig, A., et al., Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. Alzheimers Dement (Amst), 2015. 1(1): p. 112-24.
5. Tóth, L., et al. Automatic detection of mild cognitive impairment from spontaneous speech using ASR. 2015. ISCA.
6. Roark, B., et al., Spoken language derived measures for detecting mild cognitive impairment. IEEE Transactions on Audio, Speech, and Language Processing, 2011. 19(7): p. 2081-2090.
7. Moriyama, Y., et al., The Japanese version of the Rapid Dementia Screening Test is effective compared to the clock - drawing test for detecting patients with mild Alzheimer's disease. Psychogeriatrics, 2015.
8. Holtzman, R.E., et al., Social network characteristics and cognition in middle-aged and older adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 2004. 59(6): p. P278-P284.
9. Saczynski, J.S., et al., The effect of social engagement on incident dementia the Honolulu-Asia aging study. American Journal of Epidemiology, 2006. 163(5): p. 433-440.
10. Brodwin, E. New study spots patterns in President Reagan's White House talks that may have been a red flag for Alzheimer's. 2015; Available from: <http://www.businessinsider.com/changes-in-president-reagan-speech-early-sign-of-alzheimers-2015-4>. (参照 2016-06-25).
11. 池田学, 認知症. 高次脳機能研究 (旧 失語症研究), 2009. 29(2): p. 222-228.