NAIST-IS-M1651059

修士論文

自然言語処理による認知症スクリーニング

柴田 大作

2018 年 3 月 6 日

奈良先端科学技術大学院大学
情報科学研究科 情報科学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士（工学）授与の要件として提出した修士論文である．

柴田 大作

審査委員：

中村 哲 教授	（主指導教員）
松本 裕治 教授	（副指導教官）
田中 宏季 助教	（副指導教官）
荒牧 英治 特任准教授	（副指導教官）

自然言語処理による認知症スクリーニング*

柴田 大作

内容梗概

我が国における急激な高齢化に伴い，認知症高齢者の数が増加している．認知症にはその治療が困難な症状があり，早期にスクリーニングを行い，セラピー治療などにより，その進行を遅らせることが重要である．認知症のスクリーニング手法としては，PET や MRI による画像診断，Mini Mental State Examination (MMSE) による認知機能検査などが存在する. これらの手法はスクリーニングの精度は高いが，身体的侵襲や長時間の拘束が必要で，高齢者に負担を強いるものとなっている. そこで近年，言語能力に着目した認知症のスクリーニングが注目されている. 認知症になると言語能力が低下することに着目し，言語能力を測定することでスクリーニングを実現するものである. これは既存手法と比較した場合に，精度は下がるが，身体的侵襲や長時間の拘束を必要とせず，また日常で手軽に測定が可能である. 認知症では語彙量などが低下することが指摘されているが，発話の内容に踏み込んだ分析は十分になされていない．そこで本研究では発話の質的な観点を考慮した認知症のスクリーニングを行う. 質的な分析を行うために先行研究で使用された言語リソース Linguistic Inquiry and Word Count の日本語化に加え，クラウドソーシングにより人々の感情表現を収集し，分析することで大規模な日本語の感情表現辞書を開発を行った. また，英語で幅広く使用されている言語能力指標である意味密度の日本語化を行い，日本語における意味密度の算出方法を確立した. これらを用いて，発話の質的な分析を行い，認知症のスクリーニングを実現する．

キーワード

認知症，自然言語処理，言語特徴，スクリーニング，医療情報、アルツハイマー病

*奈良先端科学技術大学院大学 情報科学研究科 情報科学専攻 修士論文, NAIST-IS-MT1651059, 2018 年 3 月 6 日.

Detecting Dementia with Natural Language Processing*

Daisaku Shibata

Abstract

The number of elderly patients of dementia has been increasing in Japan because of aging society. It is important to detect and treat dementia in the early stage because treatment of dementia is difficult compared with retarding it. In order to detect dementia, diagnostic imaging (PET, MRI) and screening test for cognitive impairment (Mini Mental State Examination) have been existing. These methods are precise, though it costs huge time and money, and also requires physical invasion to patients. Methods based on natural language processing (NLP) has drawn much attention as an alternative way. To date, vocabulary size, grammatical complexity, and fluency have been studied using NLP metrics. However, the content analysis of narratives obtained from Alzheimer disease patients is still unreachable for NLP. To realize content analysis, the author translated Linguistic Inquiry and Word Count into Japanese and develop Japanese emotional dictionary via crowdsourcing. In addition to these, Idea Density, one of NLP based indexes, has been established for the detection of dementia. The aim of the study is to detect dementia with them.

Keywords:

Dementia, Natural Language Processing, Language Ability, Screening, medical informatics, Alzheimer's Disease

*Master's Thesis, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1651059, March 6, 2018.

目次

図目次

第 1 章　はじめに

1.1　背景・目的

近年，我が国における急激な高齢化に伴い認知症高齢者の数が増加している．認知症は，その治療が困難な症状（例：アルツハイマー型認知症）が存在するため，早期発見によりその進行を遅らせることが重要である．医療機関における認知症の検査方法として，PET（Positron Emission Tomography）検査や MRI（Magnetic Resonance Imaging）検査による脳画像診断や Mini Mental State Examination (MMSE) による認知機能検査などが存在する．これらの検査方法のスクリーニング精度は高いが，いくつかの問題がある．PET 検査は体内に放射性物質を注入する必要があり，身体的侵襲を伴う．そして，MRI は長時間の身体拘束が必要であり，MMSE は拘束時間は短いが，精神的な強いるものとなっている．そのため，これらの検査を日常的に受診することは難しい．

そこで近年，言語能力に基づく認知症スクリーニングが注目されている．一般的に認知症患者は健常者と比較して，言語能力が低下（使用する単語数の減少や代名詞の増加など）言語能力を自然言語処理技術を用いて，量的に測定することで認知症のスクリーニングを行うものである．言語能力による認知症のスクリーニングは，医療機関で行う診断に比べて，スクリーニングの精度は低下するが，身体的侵襲や長時間の拘束を必要としない．また，日常的に言語能力を調査することで，わずかな変化を察知し，早期発見に繋がる可能性がある．今日に至るまで言語能力に基づく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 1996 年の Nan Study[2] に始まり，最近では言語特徴に着目した研究 [5][6] や冗長性に着目した研究 [3] が報告されている．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，非常に問題を単純化してしまう恐れがある．例えば，一般的に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．そこで本研究では，発話の質的な観点を考慮した認知症スクリーニングを行う．具体的には以下に示す 3 種類の実験を行う．

1. 「不平」のような質的な概念を測定するために，先行研究で使用されている語彙を抽象化し，カテゴリ毎に分類した言語リソース の日本語化と，クラウドソーシングにより人々の感情表現を収集し，分析することで日本語の感情表現辞書を構築を行い，これらを用いた分析
2. 文章の情報量を測定する言語能力指標である意味密度の日本語における計算方法を定義し，これを用いた分析
3. 本研究で新たに算出した言語能力指標と既存の言語能力指標を用いた機械学習による認知症の識別実験

1.2 本論文の構成

第 2 章では関連研究について説明し，本研究の特徴を述べる．第 3 章では，本研究で使用するコーパスの詳細について説明する．第 4 章では，海外の言語リソース Linguistic Inquiry and Word Count (LIWC) を用いた認知症のスクリーニングについて説明する．第 5 章ではクラウドソーシングを用いた日本語の感情表現辞書の開発とそれを用いた認知症のスクリーニングについて説明する．第 6 章では日本語における意味密度の算出方法とそれを用いた認知症のスクリーニングについて説明する．第 7 章では本研究により算出した言語指標と既存の言語指標を用いて，機械学習による認知症のスクリーニングについて説明する．第 8 章はまとめとする．

第 2 章　先行研究

言語能力指標を用いた認知症のスクリーニングに関する研究は Snowdon の Nan Staudy[2] に始まった. Snowdon らは 93 人の修道女が 20 歳代の時に書いた自叙伝から，意味密度（文章中の命題数を単語数で除した値）や文法の複雑さを算出した. そして，若年期の意味密度と文法の複雑さの値が低いと，晩年期の認知機能テストの結果が低いことを示し，これらの値と認知機能の間には一定の関係があることを報告した. また，Kemper ら [4] は，文法の複雑さは認知症の有無に関わらず年齢と共に低下するが，意味密度は認知症群のみで大きく低下することを報告した.

Jarrold ら [5] は認知症者（n=22），健常者（n=23）の 15 分間のインタビュー（文献）を書き起こし，発話中の名詞，動詞などの品詞数，LIWC 頻度（詳細は第 3 章で述べる.），意味密度を抽出し，機械学習の特徴量として用いた結果，認知症者と健常者を最大 73% の精度で分類を実現したを報告した.

Orimayre ら [6] は Dementia Bank で提供されている認知症者（n=314），健常者（n=242）から構成されるコーパスからいくつかの Syntactic features と言語能力指標を抽出し，機械学習による分類を行い，74% の F 値を達成した.

荒牧ら [17] は高齢者（n=21）の「今までで一番楽しかった」ことに関する自由発話のデータを書き起こし，分析したところ，長谷川式認知症スケール（HDS-R）のスコアが低い群（n=7）で有意に潜在語彙量（話者が理解していると推定される語彙の数）が増加していることを報告した.

Sirts らは [7] 係り受け構造に基づいて意味密度の自動推定を行うソフトウェアを開発し，専門家により算出された意味密度と比較したところ強い相関（r=0.839）があり，品詞数に基づいた既存手法 [18]（r=0.795）より，精度が高いことを確認した.

さらに近年，言語特徴に加えて，音声特徴を追加して分類を行う研究が多く報告されている.

Roark ら [8] は軽度認知障害（n=37）と健常者 (n=37) から構成されるデータセットの書き起こしデータと音声データを用いて，語特徴，Syntactic features と音声特徴を抽出し機械学習による分類を行ったところ，最大で 0.861 の AUC (Area under the curve) を達成した.

田中ら [9] はコンピューターアバターとの対話を通して，言語特徴と音声特徴に

加えて，対話中のポーズ時間などの対話特徴や笑顔の割合を示す表情特徴を抽出するシステムを開発した．そして認知症者（n=14）健常者（n=15）から特徴量を抽出し，機械学習による分類を行ったところ最大で 0.93 の AUC 値が得られ，高い精度で分類できることを確認した．

Fraser ら [10] は認知症者（n=240），健常者（n=233）から構成されるコーパス（Dementia Bank）から，いくつかの言語特徴，Syntactic features と音声特徴を抽出し機械学習による分類を行ったところ，81% を超える精度での分類を達成した．これは Dementia Bank データセットを用いた研究の中では最も高い精度である．

上記のように，認知症スクリーニングに関する研究は多く存在するが，発話の内容に踏み込んだ研究は少なく，特に日本語においては，筆者の知るかぎり報告されていない．認知症者の多様な発話（例：「不安を感じる」）などの質的な部分を，先行研究で用いられている言語特徴などにより観察することは難しい．そのため本研究では，感情分析を可能にするツールを開発し，認知症者の発話内容に踏み込んだ質的な観点を考慮した複数の統計量による認知症スクリーニングの可能性について検討する．

第 3 章　実験材料

実験材料は，患者または参加者の一連の会話ごとに，1 つのコーパスを作成した．

3.1　心理検査コーパス

心理検査コーパスは，病院で診察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ものである．参加者の詳細なデータを表 3.1 に示す．

表 3.1　心理検査コーパスの詳細

No.	# of tokens	# of char	Sex	Age	MMSE
1	458	592	female	72	4
2	1945	76	female	71	14
3	1295	400	female	90	17
4	325	268	man	80	18
5	1541	2281	female	73	19
6	1177	594	female	78	19
7	1442	903	female	81	20
8	1047	1139	man	73	21
9	1796	1019	female	77	21
Mean (SD)	1225 (519)	2028 (894)	-	77 (5.7)	17 (5.0)
10	527	828	man	77	22
11	382	620	female	81	22
12	1571	2683	female	72	22
13	498	842	man	87	25
14	882	1455	man	53	25
15	1143	1934	female	87	26
16	714	1092	female	82	16
17	691	1094	man	79	28
18	920		female	71	30
Mean (SD)	814 (348)	1340 (610)	-	77 (9.9)	25 (2.6)

研究参加者のリクルートは，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 (Mild Cognitive Impairment: 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（Alazheimer's Deiseas: AD）（MMSE21 点以下）
- 対照者群 : 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（healthy Eldery: HE）（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．

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，患者に報酬は支払っていない．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．

3.2 自由発話コーパス

心理検査コーパスとは別に，本研究では中度認知症者/軽度認知症者の高齢者（以下，研究参加者）による自由会話の語りを用いた．本コーパスは研究参加者に以下の 3 つの質問を行い，その解答を録音し，書き起こしたものである．

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

本研究では図 3.2 に示すように，1 人の高齢者から 1 つもしくは 2 つの発話データを取得している．最初に 42 人の高齢者から発話データと Mini Mental State Examination (MMSE) を取得し，3 ヶ月後に継続して発話を取ることができた 32 人の高齢者から再度，発話データと MMSE の取得している．

*医の倫理委員会 承認番号: E2525

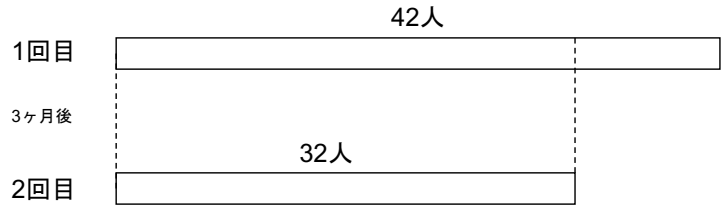

図 3.1　コーパスの構成

1 回目の参加者の詳細なデータを表 3.2 に，2 回目の参加者の詳細なデータを表 3.3 に示す.

表 3.2　1 回目の自由発話コーパスの詳細

No.	# of tokens	Sex	Age	MMSE
1	375	female	87	10
2	219	female	82	10
3	153	man	86	11
4	127	female	82	11
5	281	female	88	12
6	585	female	88	13
7	88	man	85	14
8	54	female	98	14
9	243	female	91	15
10	35	female	91	15
11	54	female	83	15
12	62	female	80	15
13	111	female	86	15
14	189	female	90	16
15	74	female	84	16
16	99	man	84	17
17	185	female	84	17
18	198	female	87	17
19	150	man	81	17
20	232	female	82	18
21	75	female	92	18
23	32	female	78	18
Mean (SD)	165 (129)	-	86 (4.5)	14.7 (2.6)
1	105	female	93	21
2	54	female	92	22
3	717	man	84	22
4	154	female	91	23
5	133	female	94	23
6	292	female	73	24
7	106	female	89	24
8	415	female	85	24
9	146	man	75	25
10	227	female	85	25
11	15	female	84	25
12	273	man	69	26
13	101	female	88	26
14	14	female	84	26
15	535	female	78	27
16	70	female	74	27
17	105	female	87	28
18	170	female	81	29
19	118	female	75	30
20	181	female	77	30
Mean (SD)	197 (178)	-	83 (7.3)	25.2 (2.9)

表 3.3　2 回目の自由発話コーパスの詳細

No.	# of tokens	Sex	MMSE	Age
1	23.0	female	7	83
2	48.0	female	9	86
3	75.0	female	10	87
4	445.0	female	11	82
5	623.0	man	12	85
6	313.0	female	13	88
7	71.0	man	14	92
8	130.0	female	15	88
9	31.0	female	16	80
10	235.0	female	17	84
11	35.0	female	18	91
12	83.0	man	18	84
13	218.0	female	18	98
14	89.0	female	19	94
15	239.0	female	19	85
16	639.0	female	19	78
17	38.0	man	19	91
18	102.0	female	20	73
19	39.0	female	20	90
Mean (SD)	198 (200)	-	86 (5.9)	15.7 (4.1)
1	476.0	female	20	81
2	253.0	female	21	82
3	415.0	man	21	84
4	464.0	female	22	91
5	94.0	man	23	84
6	66.0	female	25	84
7	110.0	female	26	84
8	84.0	female	26	75
9	39.0	female	26	93
10	77.0	female	27	78
11	96.0	female	27	75
12	1098.0	female	28	92
13	123.0	female	30	77
Mean (SD)	243 (330)	-	83 (6.3)	25.1 (2.9)

研究参加者のリクルートは，研究協力者が勤務するデイケア施設にて行った．これには以下の基準を用いた．

【包含基準】

- 医師により認知症の診断を受けている 65 歳以上の高齢者
- MMSE が 10 点以上の軽度〜中等度の認知症患者
- 「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（兵庫県但馬と京都府丹後西部，奈良県奥吉野を除く近畿地方で言語形成期を過ごしたもの）

【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
- 重度な聴覚障害者
- 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のある者

本研究の実施にあたっては「人を対象とする医学系研究に関する倫理指針」を順守し，任意参加であること，参加・不参加による不利益がないこと，いつでも参加の取り消しができることなどを対象者及び代諾者にわかりやすい言葉を用いて文書で説明した．なお，今回対象者が認知症高齢者であり，判断力の低下が中核症状として出現するため，上記指針中の「インフォームド・コンセントを与える能力を客観的に欠くと判断される成人」に該当すると判断し，本人とともに代諾者にも同意を得た場合のみに，対象者とした．なお，代諾者とは，対象者の後見人等の法的代理人，三親等以内の親族，その他これに準じる者で，対象者の最善の利益を図りうる者とした．さらに，取得したデータは，解析前に連結可能匿名化を行い，第三者がアクセスできないように，パスワード設定，施錠等で適切に管理している．研究成果の公表に際しては，個人が特定されることのないように配慮することとし，京都大学医の「医の倫理委員会」の倫理審査の承認[†]を経て実施された．

[†]医の倫理員会 承認番号：C1152

第 4 章　LIWC による認知症者の発話分析

本章では，「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 Linguistic Inquiry and Word Count（以下，LIWC）[15] 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．これまでにも認知症者 (Alzheimer's Deisease) とそうでない者 (Healthy Eldery) 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 [23] ことが一般的によく知られている．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．

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．

4.1　日本語版 LIWC の構築

4.1.1　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 4.1 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>）を作業者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 4.2 に示す．

表 4.1　LIWC2007 に収録されているカテゴリ

Total function words	Impersonal pronouns	Sadness	Inclusive
Pronoun	Article	CogMech	Excl
Personal pronouns	Common verbs	Insight	Perceptual processes
1st pers singular	Auxiliary verbs	Causation	See
1st pers plural	Past tense	Discrepancy	Hear
2nd person	Present tense	Tentative	Feel
3rd pers singular	Future tense	Certainty	Biological processes
3rd pers plural	Fillers	Inhibition	Nonfluencies
Adverbs	Family	Body	Work
Prepositions	Friends	Health	Achievement
Conjunctions	Humans	Sexual	Leisure
Negations	Affective processes	Ingestion	Home
Quantifiers	Positive emotion	Relativity	Money
Numbers	Negative emotion	Motion	Religion
Swear words	Anxiety	Space	Death
Social processes	Anger	Time	Assent

表 4.2　疾患と関連があると判断した 22 カテゴリ

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

4.1.2　LIWC の翻訳

以下の手順に沿って LIWC の日本語版を作成した.

Step 1: 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 4.2）の単語を自動的に日本語訳した.

Step 2: 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．作業者 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 4.3 に示す．

Step 3: 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．これにより，総単語数が 5,534 語から 4,769 語となった．

Step 4: 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．

Step 5: 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，作業者 1 名が目視により正しいカテゴリ（最大一つ）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 4.4）．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 4.5 に示す．各カテゴリに属する英単語と日本語の例を表 4.6 に示す．

表 4.3　誤訳の例

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

表 4.4　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

表 4.5　最終的な日本語版 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 4.6　日本語版 LIWC に収録されている単語の例

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,tirmoil	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

4.2　実験

4.2.1　手順

本実験では心理検査コーパスを用いる．以下の手順に沿って心理検査コーパスを解析した．

Step 1: 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．形態素解析には日本語形態素解析器 Janome*を使用する．

*https://github.com/mocobeta/janome

Step 2: 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．

Step 3: 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．

Step 4: Step1 から Step3 の作業を，全コーパス (n=18) に対して適用した．

4.2.2 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 4.1 に示す．また，AD と HE 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 4.7 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 4.8 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．

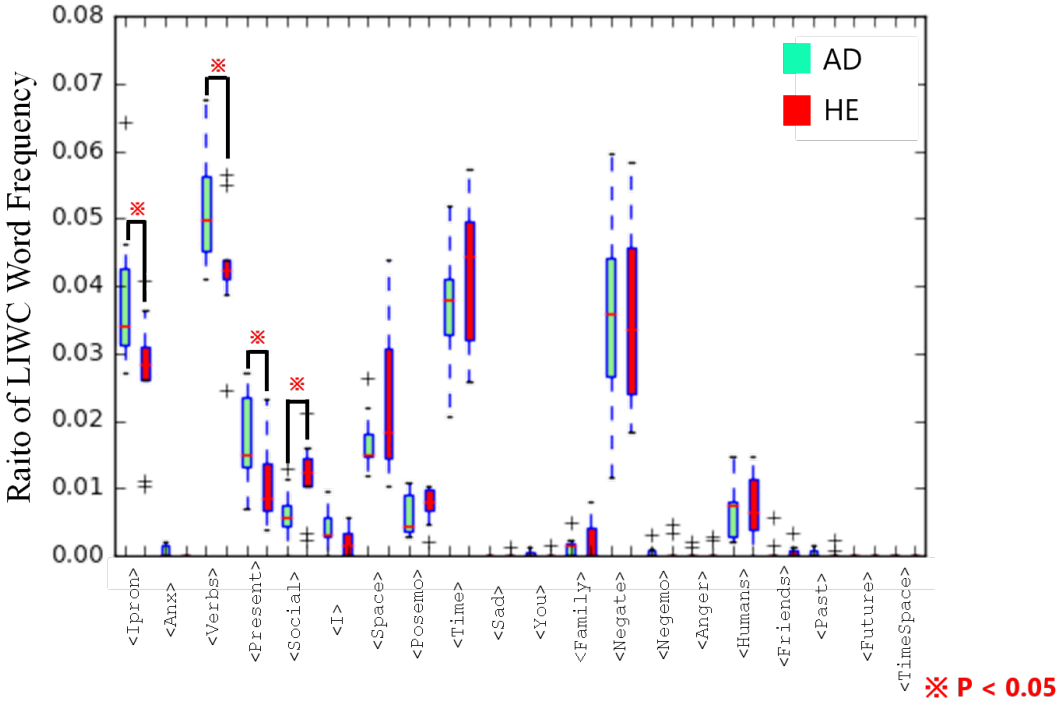

図 4.1　AD と HC の発話におけるカテゴリ使用割合のばらつき.

表 4.7　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

Category	AD	HE
Ipron	**0.0385**	**0.0268**
Anx	**0.0008**	**0.0000**
Verbs	**0.0524**	**0.0430**
Present	**0.0171**	**0.0103**
Social	**0.0063**	**0.0116**
I	0.0040	0.0019
Space	0.0170	0.0231
Posemo	0.0060	0.0076
Time	0.0364	0.0418
Sad	0.0000	0.0002
You	0.0003	0.0002
Family	0.0015	0.0021
Negate	0.0397	0.0464
Negemo	0.0006	0.0009
Anger	0.0004	0.0006
Humans	0.0068	0.0077
Friends	0.0008	0.0006
Non-Category	0.7711	0.7749
Past	0.0003	0.0003
Future	0.0000	0.0000
TimeSpace	0.0000	0.0000

表 4.8　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よ何かね
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**とかないかんから
	いうて言って**くれ**はったん

4.3　考察

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 [11]．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．

同様に，<Ipron>（非人称の代名詞）に関しては，AD で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研究で報告 [12] されている．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，不安を感じることが多くなること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 [13]．

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．

第 5 章　日本語感情表現辞書（JIWC）による認知症者の発話分析

第 4 章にて日本語版 LIWC による認知症スクリーニングについて述べたが，開発した LIWC は著作権の関係で公開することは困難であり，日本語版は未だ広く利用できる状態ではない．また，LIWC を日本語に適用するためには，LIWC もしくはコーパスのどちらかを翻訳する必要があり，これにかかるコストは大きい．さらに，LIWC には日本語には定義されないカテゴリ（冠詞）や翻訳困難な語（例えば，like という単語は<Posemo>，<Filler>，<Time>の 3 つのカテゴリに分類されている）が存在し，誤った日本語に翻訳される可能性があり，曖昧性を否定することができない．そこで，日本語のテキストに適した言語リソース，日本語感情表現辞書：Japanese Inquiry and Word Count（以下，JIWC）を構築する．具体的には，クラウドソーシングを通して不特定多数の人々のエピソードを集め，これを分析することで感情表現を抽出し，まとめた辞書を構築し，これを用いた分析を行う．

5.1　日本語感情表現辞書（JIWC）の開発

クラウドソーシングを用いた JIWC の構築手順について述べる．

5.1.1　タスク設定

Yahoo!クラウドソーシング*を用いて，不特定多数の人々が用いているリアルな感情表現を収集する．タスクとしては，表 5.1 に示すように，イギリスのコホート研究 [19] で使用された質問形式に基づいたエピソード記憶の自由記述アンケートを設定する．これは，「悲しい」出来事の回答には「悲しい」を表現する単語が含まれると考えられるためである．タスクの感情カテゴリ（以降，感情と呼ぶ）は Plutchik の感情の輪 [20]（表 5.2 に概略図を示す. また，レベルは感情の強さを表し，レベルが大きいほど強い感情である.）を参考に設定される．原則として，一般のクラウドワーカーにも馴染みがあると考えられる最も基本的な感情であるレベル

*https://crowdsourcing.yahoo.co.jp/

2 の感情表現の語を選んだが，一部質問に当てはめた際に日本語として不適切になるものは別のレベルの語に変更した（例えば，「心配」は「不安」に変更した）．今回は過去の出来事に関する質問であるため，未来に対する感情である「予測」は不要として除外する．また，「受容」は普段聞きなれない単語であり，混乱を避けるために同様の感情を意味する「信頼感」に変更する．最終的に，7 つの感情（「悲しい」「不安」「怒り」「嫌悪感」「信頼感」「驚き」「楽しい」）を用いる．また，タスクの参加者には，回答の送信をもって以下に示す回答の取扱いに同意したものとみなすことを伝えている．

- 回答内容は研究利用の範囲において，第三者に公開する可能性があること
- 調査結果を学会や専門誌で発表する可能性があること
- 回答は無記名であるため，タスクへの参加によって個人が特定されることはないこと

さらに，注意事項として回答に個人情報を含めないように促した．

表 5.1　クラウドソーシングのタスク内容

No.	質問内容
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」
9	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」

表 5.2　プルチックの感情の輪の概略図

レベル 3	レベル 2	レベル 1
悲嘆	悲しみ	感傷的
恐怖	心配	不安
苛立ち	怒り	激怒
憎悪	嫌悪感	倦怠
敬愛	信頼	容認
驚愕	驚き	動揺
恍惚	喜び	安らぎ

5.1.2　タスク投入設定

本タスクは 2017 年 4 月 3 日 17 時 00 分から 2017 年 4 月 6 日 23 時 40 分にわたって実施され，1,000 名のクラウドワーカーが参加した．

5.1.3　タスクの回答結果

本タスクによって取得された回答結果の一例を表 5.3 に示す．各感情の No. は表 5.2 と対応している．

回答結果に以下の内容を含むものは，作業者 2 人が目視により削除を行なった．

- 差別的な内容を含む回答
- 政治的，宗教的な内容を含む回答
- 個人情報を含む回答
- 攻撃的な単語を含む回答

表 5.3　回答結果の例

No.1: 悲しい	両親が立て続けに亡くなったこと 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
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たとき，天候にイカリ
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
No.5: 信頼感	いい人と結婚できたこと 困ってい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
No.6: 驚き	子供が受験に合格したこと 大学に合格した時
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

5.2　日本語感情表現辞書（JIWC）の構築

3.2 節で取得したデータを用いて，JIWC の構築を行う．

Step 1: 形態素解析

回答結果の形態素を取得する．形態素解析器は日本語形態素解析器（Juman++）[†]を使用し，抽出する形態素は品詞が名詞，動詞，形容詞と副詞であるものに限定する．

Step 2: 単語頻度・割合の算出

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する．この際，各単語の合計頻度が閾値以下の単語は削除する．各感情の単語頻度を各単語の合計頻度で除すことで単語割合を算出

[†]http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN

する．

Step 3: ノイズ除去

Step2 において，単語の出現頻度が 10 以下の単語については削除を行う．

Step 4: TF-IDF による重み付け

TF-IDF は単語頻度を示す TF (Term Frequency) と逆文章頻度を示す IDF (Inverse Document Frequency) の積で算出される．TF-IDF の定義を以下に示す．

Step 5: 分類

単語ごとに TF-IDF 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加える．なお，単語スコアの最大値が複数の感情で同じ場合は，1 つの単語が複数の感情に属することとする．

$$tf - idf(w_i, D_j) = tf(w_i, D_j) \times idf(w_i, D_j)$$

$$idf(w_i, D_j) = \log \frac{|D_j|}{1 + tf(w_i, D_j)}$$

w_i: i 番目の単語.

D_j: j 番目のエピソード.

$tf(w_i, D_j)$: j 番目のエピソード D_j における単語 w_iの出現頻度

$idf(w_i, D_j)$: ある単語 w_iの逆文章頻度

構築した日本語感情表現辞書に収録されている単語の例を表 5.4 に示す.

表 5.4　日本語感情表現辞書に含まれる単語の例

悲しい	不安	怒り	嫌悪感	信頼感	驚き	楽しい
がる	いける	お金	あげる	あと	いた事	いく
しまう	うまい	せる	あげる	ある	する	すぎる
ずっと	かかる	つかれる	あまり	いつも	とくに	てる
なくなる	こども	られる	いい	いろいろ	ところ	とても
ばあちゃん	これから	れる	いう	かける	ぶり	みんな
まだ	どう	トラブル	いじめ	くれる	よく	やる
もう	なかなか	ネット	いる	これ	中学生	アーティスト
ペット	なる	ミス	くる	さん	予想	ゲーム
ヶ月	わかる	事件	こちら	すぐ	以上	テレビ
事故	バイト	他人	すべて	すごい	会う	デート
亡くす	パニック	会社	それ	ため	偶然	ハワイ
亡くなる	リストラ	使う	たくさん	できる	再会	バイク
交通	上手い	働く	つく	もらう	出る	パチンコ
他界	不安	全く	でる	もらえる	出産	ライブ
付き合う	今後	勝手	ない	パートナー	勉強	一番
元気	仕方	北朝鮮	みる	一生懸命	取る	一緒
出来事	住む	受ける	もつ	両親	受かる	中学
別れ	体調	問題	ゴキブリ	主人	合格	乗る
別れる	借金	喧嘩	トイレ	仕事	同級生	人生
大事	健康	嫌がらせ	ニュース	信頼	告白	仲間
大切	入試	対応	マナー	先生	地元	会話
大好き	入院	彼氏	上司	先輩	場所	全て
失恋	分かる	後輩	不倫	出会い	多い	出かける
失敗	卒業	怒り	人間	出来る	大きい	初めて
小さい	原因	怒る	介護	助ける	妊娠	夫婦
小学生	受験	怒鳴る	付き合い	友人	嬉しい	好き
当時	地震	扱い	信じる	友達	宝くじ	子ども
彼女	大学	担当	入る	同僚	実家	子供
恋人	大震災	政治	出す	困る	居る	子育て
悲しい	始める	教師	出会う	変わる	当たる	学生
愛犬	学校	欲しい	合う	存在	当選	家庭
振る	将来	浮気	周り	守る	思い	家族
暮らす	就職	無視	大人	帰る	思う	忘れる
最愛	常に	理不尽	女性	強い	息子	思い出
最近	年金	盗む	嫌い	得る	懸賞	思い出す
死ぬ	待つ	社員	嫌悪	必ず	成長	新婚
死別	手術	社長	悪い	感じる	日本	旅行
死去	新しい	約束	悪口	感情	早い	日々
泣く	時期	職場	態度	持つ	普通	時代
父親	東日本	自転車	抱く	支える	治る	時間
祖母	決まる	色々	日常	旦那	特に	本当に
祖父	活動	裏切り	生理	普段	生まれる	東京
祖父母	状態	裏切る	男性	来る	発覚	楽しい
経験	現在	覚える	痴漢	母親	知り合い	毎日
自殺	生きる	言う	発言	気持ち	知る	海外
若い	病気	詐欺	知人	決める	突然	生活
落ちる	病院	貸す	経営	無い	精神	笑う
親族	発表	返す	行動	理解	結婚	結婚式
赤ちゃん	社会	部下	行為	相手	聞く	考える

5.3 実験

5.3.1 分析

2 章で述べた 2 種類のコーパスに対して，以下の方法により分析を行う．

Step 1: コーパスの形態素解析コーパスごとに形態素解析を行い，コーパス中の代表表記を取得する．形態素解析には節 5.2 と同様に日本語形態素解析器（Juman++）を使用する．

Step 2: JIWC による分析取得した代表表記と JIWC をマッチさせ，各感情の使用頻度を算出する．具体的には Step1 で取得した代表表記が各感情の感情表現と一致するかをチェックする．これにより各感情の使用頻度を算出する．そして，コーパスごとの使用頻度を コーパスごとの全代表表記数で除し，使用割合に変換した．

5.3.2 心理検査コーパス分析結果

心理検査コーパスから算出した各感情の使用割合について AD 群と HE 群で t 検定による比較を行った．その結果を表 5.4 に示す．表 5.4 から，「不安」，「怒り」，「嫌悪感」に分類された感情表現の使用割合が AD 群で有意に増加していることがわかる．「悲しい」，「信頼感」，「驚き」，「楽しい」については有意差が確認されなかった．

表 5.5　心理検査コーパスの分析結果．有意差が確認された項目を*で示す．

	AD 群平均値	HE 群平均値
悲しい	0.037	0.054
不安	0.087*	0.059*
怒り	0.129*	0.101*
嫌悪感	0.103*	0.060*
信頼感	0.128	0.110
驚き	0.058	0.059
楽しい	0.130	0.071

5.3.3　自由会話コーパス分析結果

自由会話から算出した各感情の使用割合について中度認知症群と軽度認知症群で t 検定による比較を行った．その結果を表 5.5 に示す．表 5.5 から，「不安」に分類された感情表現の使用割合が中度認知症群で有意に増加していることがわかる．「悲しい」，「怒り」，「嫌悪感」，「信頼感」，「驚き」，「楽しい」については有意差が確認されなかった．

表 5.6　自由会話コーパスの分析結果．有意差が確認された項目を*で示す．

	中度認知症群平均値	軽度認知症群平均値
悲しい	0.070	0.063
不安	0.081*	0.052*
怒り	0.129	0.125
嫌悪感	0.103	0.093
信頼感	0.128	0.147
驚き	0.058	0.063
楽しい	0.130	0.146

5.4　考察

5.4.1　心理検査コーパス分析結果に関する考察

どのような理由でいくつかのカテゴリにおいて有意差が見られたのであろうか．認知症では，その症状の進展に伴う様々な日常生活における失敗から怒りやすくなることや，不安を感じることが一般的に言われている．「怒り」と「不安」が有意に増加していることは，これらの特徴を定性的に評価している可能性があり今後インタビューなどを通して定性的に調査したい．一方，感情「嫌悪感」については現時点で十分な考察ができていないため，今後の課題としたい．

5.4.2　自由会話コーパス分析結果に関する考察

どのような理由で「不安」が中度認知症群において有意に増加したのか．これは協力施設内で認知症者の方に自由に喋ってもらうというスタイルで行ったため，施設で普段行わないイベントが発生し，不安を覚えたためということが考えられる．

5.4.3　全体的な考察

心理検査コーパスと自由会話コーパス，両コーパスにおいて「不安」が AD 群と中度認知症群の両群で有意に増加していることが確認された．認知症ではその症状（記憶の混乱など）から不安を感じるということが一般的に言われており，語りの中に隠れている「不安」がこのように数値に現れたのではないかと考えられる．我々が開発した日本語感情表現辞書（JIWC）を用いることで，直接的に感じることが難しい語りの中に含まれる感情を定量的に評価することの可能性が示唆された．本研究で調査対象としている高齢者（平均年齢：80.6 歳）とクラウドワーカーの年齢層にギャップがあり，使用する感情表現が異なる可能性がある．これについては今後，年齢ごとに層別化した辞書を開発することで解決することを考えている．

第 6 章　日本語における意味密度の導入

認知症と意味密度には強い関係性があることが報告されている．また，認知症を識別する際の特徴量として，いくつかの研究で使用されている．英語では幅広く使用されている言語指標であるが，この日本語版は長らく存在してこなかった．そこで，意味密度のアノテーションのルール [16] を英語から日本語に翻訳し，日本語における意味密度の算出方法を定義する．また，人手による意味密度の算出はコストが大きいことから，日本語における意味密度の簡易推定法を提案し，これの有用性と認知症との関係について調査する．

6.1　意味密度のとは

意味密度は言語を扱う心理学者によって定義 [24] されたものであり，文章の意味内容の豊かさを示す指標である．先行研究において若年期の意味密度晩年期の認知症の発症リスクには一定の関係がああることが報告されている．一般的に意味密度は文章中の命題数を単語数で除した値で定義され，命題は True もしくは False で答えることができるものである．

例えば，以下に示す"The old gray mare has a very large nose."という文章には (HAS，MARE，NOSE)，(OLD，MARE)，(GRAY，MARE)，(LARGE，NOSE) と (VERY，(LARGE，NOSE)) という 5 つの命題から構成され. 命題数が 5，単語数が 9 であることから意味密度は 0.56 となる．

(1) a. The old gray mare has a very large nose.
    1. HAS，MARE，NOSE
    2. OLD，MARE
    3. GRAY，MARE
    4. LARGE，NOSE
    5. VERY，(LARGE，NOSE)

6.2 日本語における意味密度の算出

英語における意味密度の算出方法を日本語に適用し，日本語における意味密度の算出を行う．意味密度の算出方法は文章中の命題数を単語数で除した値と定義されており，日本語における単語数と命題数の算出方法について 6.2.1 節と 6.2.2 節に示す．

6.2.1 単語数

単語数は先行研究 [16] に基づき，日本語形態素解析器 (Juman++) によりコーパスの形態素解析を行い，品詞が名詞，形容詞，連体詞，形容動詞，動詞，副詞，接続詞，感動詞である形態素数とする．

単語数を上記の品詞である形態素に限定した理由は，日本語の文法的特徴による．日本語の文章を形態素解析すると，助詞や助動詞，接辞などの機能語も含む全ての形態素が分割されてしまい，形態素の総数はかなりの数になる．したがって，単純に命題数を総形態素数で除して意味密度を算出することは，総単語数をベースにしている先行研究（英語）との対応を考えると適切ではない．そこで，先行研究 [16] を参考に，日本語で命題を構成すると考えられる上記の品詞に限定して形態素数を規定した．

6.2.2 命題数

命題のアノテーション方法は，先行研究 [16] の規定 Analysis of Idea Density (AID) に基づいて構築した．AID では，述語 (Predication)，修飾語 (Modification)，接続詞 (Connectives) が命題の主要部を構成する要素とされており，本研究もこれに従った．しかし，AID は英語のアノテーション方法についてのマニュアルであり，全てを日本語に適用することは困難である．そのため適用が困難な規則や日本語特有の事情に関しては，言語研究者の判断で適宜変更を加えた．変更内容は以下の通りである．

述語

日本語において，述語を構成する品詞は，（Juman++ の品詞体系における）動詞，名詞，形容詞，形容動詞である．アノテーションに際しては，述語およびその項をひとまとまりとして，一つの命題とみなした．「述語，主語，（目的語）」の順で表記する (2)．この時，時制，アスペクト，モダリティを表す要素については，命題の数に影響しない要素とみなした．また，日本語では主語が明示されないことが多い．そのため，命題を表記する際には主語をカッコ書きで補って表記した (3)．なお，この際補った主語はアノテーションの都合上付されているだけであり，意味密度の算出に際しての形態素数の計上とは無関係である．

以下は，動詞のアノテーション例である．

(2) a. 私は何でも食べます．
    1. 食べます，私は，何でも

(3) a. 何でも食べます．
    1. 食べます，（私が），何でも

なお，補助動詞は独立した動詞とは認めず，補助動詞に前接する動詞についてアノテーションをする．

(4) a. 兄にご飯を作ってもらった．
    1. 作ってもらった，兄に，ご飯を

形容詞の叙述用法は，動詞と同様に表記する．

(5) a. その車は青い．
    1. 青い，その車は

形容動詞も形容詞と同様にアノテーションを行う．

(6) a. そこは静かだった．
    1. 静かだった，WHERE
    2. WHERE ＝ そこは

修飾語

日本語における修飾語は，形容詞，形容動詞，副詞，連体詞である．この他に，時・場所を表す名詞句，否定を独立した命題とみなす．

形容詞・形容動詞は，限定用法の場合 (7, 8)，独立した命題とみなすが，前節に示した通り叙述用法の場合 (5, 6) は動詞と同様に扱う．

(7) a. 青い車を買った．
    1. 買った，(私が)，車を
    2. 車，青い

(8) a. 静かな車を買った．
    1. 買った，(私が)，車を
    2. 車，静かな

副詞は，独立した命題とみなす．

(9) a. ラーメンをよく食べました．
    1. 食べました，(私が)，ラーメンを
    2. 食べました．よく

時・場所を表す名詞句は，それぞれ独立した命題とみなす．

(10) a. 昨日映画を見た．
    1. 見た，(私が)，映画を
    2. WHEN = 昨日

(11) a. 学校に行く．
    1. 行く，(私が)，WHERE
    2. WHERE = 学校

ただし，時や場所が，主語が明らかに復元できる動詞を含む従属節によって表されている場合，その従属節についても命題を計上する．

(12) a. 急いで学校に行く途中で転んだ．

1. 行く途中で，（私が），学校に
2. 行く，急いで
3. 転んだ，（私が）

否定は，独立した命題とみなす．この時，動詞句の命題は肯定形で表記し，さらに否定辞を別の命題として表記する．

(13) a. 私はご飯を食べなかった．
1. 食べる，私は，ご飯を
2. NEG なかった

6.3 実験 1: 機械翻訳による意味密度の算出

英語における意味密度の算出手法を，機械翻訳を用いることで日本語に対して適用する．

6.3.1 実験方法

英語における意味密度の自動計算手法

先行研究において，英語テキストにおける意味密度の自動計算手法が 2 つ提案されている．1 つは Computerized Propositional Idea Density Rater (CPIDR)[18]，もう一方は dependency-based Propositional idea density (DEPID)[7] である．詳細を以下に示す．

- CPIDR
  テキスト中の動詞，形容詞，形容動詞，前置詞と等位接続詞の数を算出し，それを 37 個のルールを適用することで意味密度を計算する．
- DEPID
  いくつかの係り受け構造は命題と共通していることに着目したもので，テキスト中の命題に相当する係り受け数を算出し，意味密度を計算する．

機械翻訳

章 6 からわかるように，人手による命題のアノテーションはコストが大きく，現実的ではない．しかし，英語テキストであれば先行研究で使用されている意味密度の算出方法 (CPIDR や DEPID など) を適用することができる．そこで、コーパスを機械翻訳を用いて英語に翻訳（翻訳コーパス）し，6.2.1 節の手法を適用することで意味密度の算出を行う．機械翻訳にはインターネット上で利用可能な機械翻訳エンジン（表 6.1）を使用する．

本研究では，DEPID は CPIDR より正確に命題数を算出できると報告されていることから，DEPID を用いて意味密度の算出を行う．機械翻訳を用いた意味密度は先行研究 [7] に基づき，DEPID により算出した翻訳コーパス中の命題数を，翻訳コーパス中の総単語数で除した値とする．

$$IdeaDensity = \frac{NumberofIdeas}{Numberoftokens} \tag{6.3.1}$$

表 6.1　機械翻訳一覧

サービス	手法	提供
Google 翻訳	ニューラルトランスレーション	Google
Bing 翻訳	統計的機械翻訳	Microsoft
excite 翻訳	フレーズベース機械翻訳	Excite Japan

比較

機械翻訳によりテキストを翻訳することで，命題数に影響を与える可能性がある．そのため，日本語において算出した命題数と機械翻訳により算出した命題数の関係性の確認を行う．

6.3.2　結果と考察

日本語のテキストについて人手により算出した命題数と，翻訳テキストについて DEPID により算出した命題数の相関係数と傾きを図 6.1 と表 6.2 に示す．相関係

数は Google 翻訳を用いたものが最も高く（r = 0.983），傾きは Bing 翻訳を用いたものが最も 1 に近い（$\alpha = 0.913$）ことが確認された．

機械翻訳で算出した命題数は，日本語で算出した命題数と比較して，最低でも 10% 程度の誤差が確認された．先行研究 [18] と比較して約 2 倍程度の誤差であるが，本研究ではテキストの翻訳を行なっているため，単縦な比較は困難である．しかし，両者の間には強い相関が確認されており，簡易推定という点では十分な精度であると考えられる．

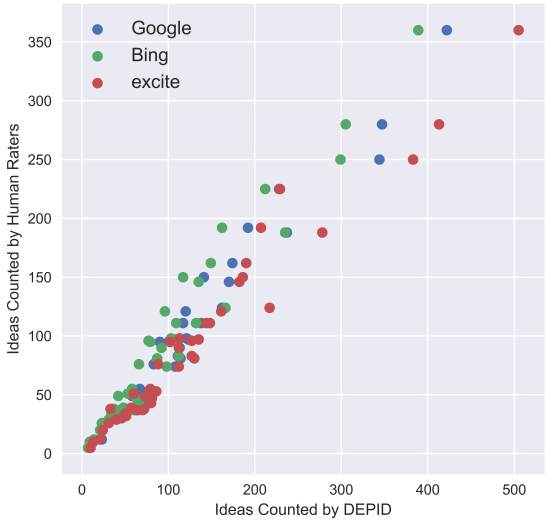

図 6.1　日本語で算出した命題数と機械翻訳で算出した命題数の相関

表 6.2　相関係数による評価

翻訳方法	相関係数	傾き
Google	0.983	0.838
Bing	0.975	0.913
excite	0.978	0.714

6.4 実験 2: 意味密度による認知症者の発話分析

意味密度の認知症スクリーニングへの適用の可能性について検討するため，人手により算出した意味密度，機械翻訳により算出した意味密度と MMSE の相関性について調査を行う．

6.4.1 算出方法

意味密度の算出は DEPID により行うが，DEPID の改良型として，DEPID-R と DEPID-R-ADD があり，これらを含めた意味密度の算出を行う．DEPID-R と DEPID-R-ADD の特徴を以下に示す．

- DEPID-R
  認知症者の発話の特徴として，同じ内容を繰り返すことがある．同じ内容を繰り返すと命題数が増えてしまい，意味密度が高くなる．そこで，全ての命題数ではなく，延べ命題数を用いて意味密度の算出を行うものが DEPID-R である．DEPID-R により算出した命題数を DR-Idea と表記し，以下に意味密度の算出方法を示す．

$$IdeaDensity = \frac{No.ofDR - Ideas}{No.oftokens} \tag{6.4.1}$$

- DEPID-R-ADD
  DEPID-R に曖昧な文章に含まれる命題をカウントしない機能を追加したものが DEPID-R-ADD である．認知症者は，具体的な命題を含まない曖昧な文章を多く発話する特徴がある．
  そのため speciteller[21] により文章の具体性を算出し，具体性が低い曖昧な文章を除外する．先行研究と同様に Speciteller's specificity rating が 0.01 以下の文章に含まれる命題はカウントしない．DEPID-R-ADD により算出した命題数を DRA-Idea と表記し，以下に意味密度の算出方法を示す．

$$IdeaDensity = \frac{No.ofDRA - Idea}{No.oftokens} \tag{6.4.2}$$

日本語で算出した意味密度を J-ID とする．また，DEPID，DEPID-R，DEPID-R-ADD により算出した意味密度をそれぞれ D-ID，DR-ID，DRA-ID とする．

6.4.2 J-ID と MMSE

J-ID と認知機能スコア (MMSE) との相関係数は-0.039 (図 6.2) と，相関は確認されなかった．

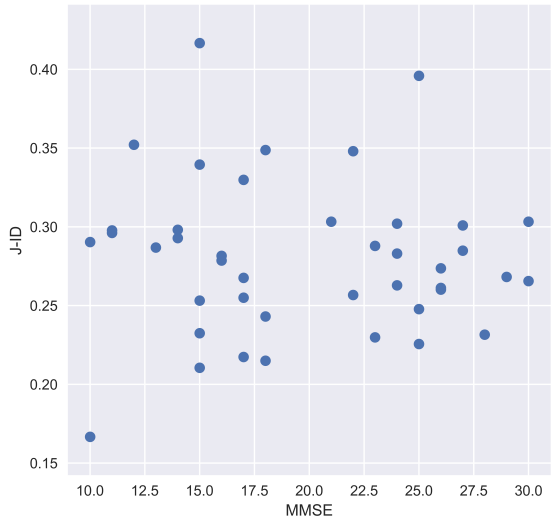

図 6.2　日本語における意味密度と認知機能スコアの相関

6.4.3 D-ID と MMSE

D-ID と MMSE の相関は最大で 0.290(表 6.3) であり，弱い相関が確認された．

表 6.3　相関係数による評価

翻訳方法	MMSE との相関係数
Google	0.100
Bing	0.120
excite	0.290

6.4.4 DR-ID · DRA-ID と MMSE

DR-ID と MMSE の相関は最大で 0.386（表 6.4）と弱い相関が確認され, DRA-ID と MMSE の相関は最大で 0.473（表 6.4）と中程度の相関が確認された.

表 6.4 相関係数による評価

手法	翻訳方法	相関係数
DEPID-R	Google	0.160
	Bing	0.060
	excite	0.386
DEPID-R-ADD	Google	0.473
	Bing	0.334
	excite	0.365

6.4.5 結果

J-ID と MMSE の相関は確認されなかった（r= 0.100）. D-ID と MMSE ならびに DR-ID と MMSE では，翻訳方法に excite 翻訳を用いた際に，それぞれ弱い相関（r = 0.290, 0.386）が確認された．また，DRA-ID と MMSE では，Google 翻訳を用いた際に中程度の相関（r=0.473）が確認された.

6.4.6 考察

日本語で算出した意味密度と MMSE に相関は確認されなかった. しかし，機械翻訳を用いて算出した意味密度と MMSE の間には最低でも弱い相関が確認されており，日本語と英語の命題数には強い相関関係があることから，問題点は単語数の算出方法である可能性がある．意味密度を算出する際，日本語では命題数を品詞を限定した形態素数，英語では総語数で除しているが，この 2 つは異なるものであり，この差が相関に影響を与えていることが考えられる．そのため，日本語において，英語の単語に相当するものを定義する必要がある.

Google 翻訳により翻訳したテキストを DEPID-R-ADD (重複する命題と曖昧な文章を除外したテキストから命題数を算出する方法) に適用し，算出した意味密度と MMSE の相関係数は 0.473 と中程度の相関が確認された．英語テキストで曖昧な文章が，認知症をスクリーニングするための意味密度という点でノイズになることは報告されているが，日本語から機械翻訳により英語にしたテキストにおいても同様の結果が得られた．これは英語テキストと同じ手法が機械翻訳を用いることで認知症スクリーニングに適用できる可能性を示しており，新たな知見である．

この理由として，日本語の発話の特徴として主語の省略があることから，日本語で意味密度を算出することは手動，自動に関わらず困難である．具体的には人手による命題のアノテーションにおいて，主語が省略されている場合は，主語を補う必要がある．例えば，「映画を見た」という文章に対してアノテーションを行うと以下のようになり，主語を補っていることがわかる．

(14) a. 映画を見た．
　　1. 見た，私が，映画を

しかし，英語に翻訳することで文章の主語が補われることがある．そのため，機械翻訳により日本語テキストを英語に翻訳し，算出する手法は日本語における意味密度を自動で算出するにあたり，有効な手法となっている可能性がある．

第 7 章　機械学習による軽度認知症と中度認知症の識別

本章では，機械学習を用いた軽度認知症者と中度認知症者の自動識別の可能性について検討する．

本研究では，先行研究 [25][26] に基づき，自由発話コーパスにおける MMSE が 21 点以上の参加者を軽度認知症，20 点以下の参加者を中度認知症とする．

7.1　特徴量の抽出

自由発話コーパスに対して，日本語形態素解析器（Juman++）を用いた形態素解析を行い，品詞が名詞，動詞，形容詞と副詞である単語を取得し，それに基づいて以下に示す言語特徴の作成，抽出を行った．

本研究では，7.1.1 節と 7.1.2 節の特徴を単語頻度特徴，7.1.3 節から 7.1.6 節までを言語特徴と定義する．

7.1.1　Bag of Words 特徴（BOW）

BOW はコーパスの全体集合から，単語から成り立つ語彙を作成し，各コーパスでの各単語の出現頻度を含んだ特徴ベクトルを構築するモデルである [22].

7.1.2　LIWC 特徴 (LIWC)

LIWC 特徴は，4 章において述べた日本語版 LIWC を用いて抽出を行う．抽出の手順は 4.2.1 節と同様の手順で行い，コーパス中の各カテゴリの使用割合を特徴量とする．

7.1.3　JIWC 特徴 (JIWC)

JIWC 特徴は，5 章において述べた JIWC 用いて抽出を行う．抽出の手順は 5.3.1 節と同様の手順で行い，コーパス中の各感情の使用割合を特徴量とする．

7.1.4 言語能力特徴（Language ability scores: LAS）

言語特徴として，以下に示す 7 つの特徴量の抽出を行う．

- Token
  コーパス中の異なり単語数を示す．
- Type
  コーパス中の延べ単語数を示す．
- Type Token Ratio (TTR)
  Type と Token の比率（Type/Token）で算出される．一般的に，この値が大きいほど語彙量が大きいとされている．
- 平均文字数
  単語の平均文字数を示す．
- 語彙難易度 (Japanese Educational Level: JEL)
  語彙の難易度を示す指標である．難易度は日本語学習辞書*に収載されている語彙レベルを用いた．語彙レベルは，1(初級前半)，2(初級後半)，3(中級前半)， 4(中級後半)，5(上級前半)，6(上級後半) に分けられる ．語ごとに算出し，全単語の平均を JEL スコアとする．
- 潜在語彙量 (Potential Vocabulary Size: PVS)
  潜在語彙量は，先行研究 [17] によって示されるように，ある人物が無限時間発話した際に使用する可能性のある語彙量の予測である．対象者の用いた延べ単語数を計測し，それにジップ則を用いることで推測される．

7.1.5 意味密度（ID）

意味密度とは，話者が発話した文がどの程度豊かな意味内容を伴っているかということを測定するための指標である．意味密度の計算方法は，通常，1 つの文章に含まれる命題数を文章の異なり単語数で除した値とされる．本実験では，6.2.1 節，6.3.1 節と同様に，係り受け構造に基づいた 3 種類の意味密度 (D-ID，DR-ID，

*http://jhlee.sakura.ne.jp/JEV.html

DRA-ID) を算出し，特徴量として用いる．

7.2　実験

実験として節 7.2.1 と 7.2.2 に示す 2 種類の実験を行う．

7.2.1　実験 1: 交差検証による評価実験

自由発話コーパスにおいて，1 回目と 2 回目で発話データを取得することができた 32 人，64 つの発話データ（図 7.1）を用いた交差検証による評価実験を行う．

機械学習のモデルには先行研究で使用されているロジスティック回帰を使用する．BOW， LAS， ID, LIWC, JIWC, LAS + ID, LAS + LIWC, LAS + ID + LIWC, と LAS + ID + LIWC + JIWC を入力し，学習を行った 10 個のモデルを作成した．

モデルの評価は 5-fold 交差検証を用いて，ROC（Receiver operator Characteristic）曲線を描画し，AUC（Area Under the Curve）を算出することで行う．

ROC 曲線はモデルの診断性能の特性を表現する方法で縦軸に偽陽性，横軸に真陽性の割合をプロットすることで描画される．モデルの検出がランダムである場合には 45 度の直線になり，診断性能が高いほど弓なりの曲線となる．この曲線下の面積の割合が AUC であり，診断性能が高いほど 1 に近い値となる．

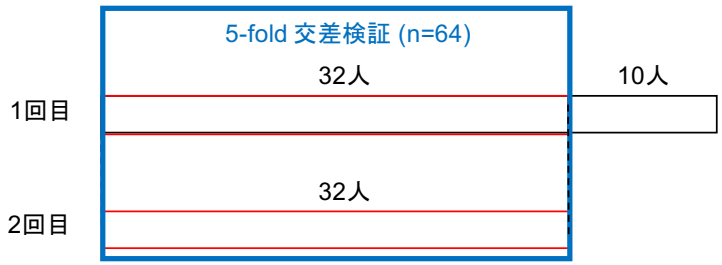

図 7.1　実験 1: 交差検証による評価実験

7.2.2　実験 2

自由発話コーパスにおいて，1 回目と 2 回目で発話データを取得することができた 32 人，64 つの発話データを学習データとし，1 回目でのみ発話を取得することができた 10 人の発話データをテストデータ（図 7.2）とした評価実験を行う．作成するモデルは節 7.2.1 と同様で，モデルの評価も同様に ROC 曲線を描画し，AUC を算出することで行う．

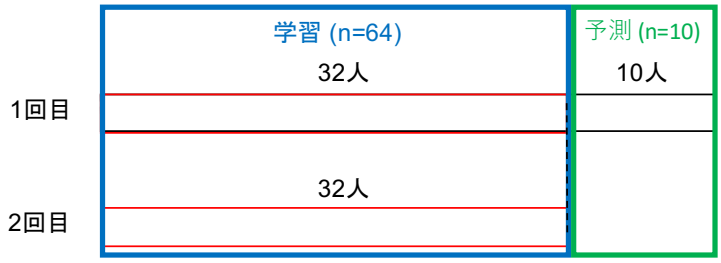

図 7.2　実験 2:テストデータによる評価実験

7.3　実験結果

7.3.1　実験 1 の結果

実験 1 の結果を表 7.1 に示す．入力の特徴量として，LAS と LIWC を組み合わせて学習を行なったモデルの AUC が 0.710 と最も大きいことが確認された．

表 7.1　実験 1 の結果

No.	Features	AUC
1	LAS + ID + LIWC + JIWC	0.690
2	LAS + ID + LIWC	0.700
3	LAS + LIWC	**0.710**
4	LAS + ID	0.556
5	ID + LIWC	0.596
6	LIWC	0.626
7	LAS	0.622
8	ID	0.562
9	JIWC	0.593
10	BoW	0.629

7.3.2　実験 2 の結果

実験 2 の結果を表 7.2 に示す. 入力の特徴量として，LAS + LIWC, LAS + ID + LIWC と LIWC + ID + LIWC + JIWC を組み合わせて学習を行なったモデルの AUC がいずれも 1.000 と最も大きいことが確認された.

表 7.2　実験 1 の結果

No.	Features	AUC
1	LAS + ID + LIWC + JIWC	**1.000**
2	LAS + ID + LIWC	**1.000**
3	LAS + LIWC	**1.000**
4	LAS + ID	0.857
5	ID + LIWC	0.905
6	LIWC	0.905
7	LAS	0.762
8	ID	0.762
9	JIWC	0.619
10	BoW	0.333

7.4　考察・まとめ

実験 1 において，AUC が最大で 0.71 と構築したモデルは中程度のスクリーニング性能を有することが確認された．得られた結果と先行研究を比較することは，使用しているコーパスや特徴量が異なるため困難であるが，言語特徴だけを用いた簡易的なスクリーニングという観点においては十分な性能であることが考えられる．

実験 2 において，AUC が最大で 1.00 と構築したモデルは高いのスクリーニング性能を有することが確認された．この 1.00 という値はある閾値において完全な識別が可能であることを示している．しかしテストデータの数が少なく（n=10），偶然起こりうる可能性があり，信頼性に欠ける．そのため今後，コーパスサイズを大きくし，改めて検討することが必要である．

モデルの入力として，単語頻度特徴を用いたモデルより，言語特徴を用いたモデルや組み合わせたモデルの方が AUC が高いことが確認された．一般的に文章分類のタスクでは BOW などの単語頻度が入力として利用されるが，認知症のスクリーニングでは言語特徴を組み合わせて用いた方が精度が得られる可能性がある．また，LAS と LIWC を組み合わせた場合の精度が最も高かったことから，日本語のスクリーニングにおいて語彙量や品詞割合などの言語能力特徴だけでなく，内容に踏み込むことができる特徴も重要である可能性が示唆された．

ID と JIWC の有効性は実験結果からは確認することはできなかった．ID は健常者と認知症者において差が確認されている特徴であり，本研究では軽度，中度認知症者を対象としており疾患の違いが影響していると考えられる．

JIWC は，クラウドソーシングを用いて開発を行なったが，クラウドワーカーの平均年齢は 40 代であることが確認された．本実験の参加者の平均年齢は 84.5 歳であり，平均年齢に差がある．そのため，使用する言葉が異なり，それが影響した可能性がある．

第 8 章　総括

本研究では，発話の質的な分析による認知症スクリーニングに向けて，自然言語処理を用いた認知症者の発話の質的な分析を行なった．具体的には海外の言語リソースである Linguistic Inquiry and Word Count (LIWC) の日本語化（第 4 章），日本語感情辞書（JIWC）の構築（第 5 章），日本語における意味密度の定義とこれらを用いた認知症者の発話分析（第 6 章），加えて機械学習による認知症のスクリーニング（第 7 章）に関する検討を行なった．LIWC を用いた認知症のスクリーニングでは，認知症者は健常者と比較して，代名詞や不安を表す単語の使用割合が有意に増加することが確認された．認知症者の特徴として，一般的に言われている「代名詞の増加」や「周辺症状により不安を感じやすくなる」などを定量的に評価することができた．JIWC を用いた認知症のスクリーニングでは，これまで開発されてこなかった日本語の感情表現辞書をクラウドソーシングを用いて開発し，軽度認知症者と中度認知症者の発話の分析を行い，中度認知症群において，「不安」を表す単語の使用割合が有意に増加していることが確認された．認知症の程度に比例して，「不安を感じやすくなる」といったような特徴は強くなる可能性が示唆された．また，先行研究においてその有用性が示されている意味密度の日本語における定義，簡易の推定方法と意味密度と認知症の進行具合との関係性を調査した．提案した手法を用いて算出した意味密度と人手により算出した意味密度には強い相関（r=0.983）があり，MMSE との相関は最大手 0.473 と中程度の相関が確認され，提案手法の有用性を示すことができた．そして機械学習を用いた軽度認知症者と中度認知症者の識別実験では，本研究において算出した言語指標と既存の言語指標を組み合わせて学習することで，AUC が最大 0.710 であることが確認され，中程度のスクリーニング性能を有するモデルを構築できることがわかった．

今後の課題としては，コーパスサイズを大きくし，より正確な分析とモデルの構築を行うことが挙げられる．本研究で使用したコーパスサイズは，海外の先行研究で使用されているコーパスサイズの約 1 割程度である．また，言語能力を測定するシステムを開発し，言語能力からの認知症スクリーニングが実際に可能なのかを今後，検討する必要があると思われる．

謝辞

本研究の副指導教官である荒牧英治特任准教授には，熱心なご指導及びご助言を賜りました. 日頃より，研究活動を進めやすいようにと，研究環境を整えて下さったことに心より感謝申し上げます．また，主指導教官である中村哲教授, 副指導教官である松本裕治教授, 田中宏季助教には，御多忙の中にも関わらず，本研究の論文審査委員を引き受けて頂き，心より感謝申し上げます. また，中間発表の際には重要なご意見をくださり，自らの研究を見直す契機となりました. 重ねて御礼申し上げます．そして，伊藤薫研究員と若宮翔子博士研究員には，研究テーマの立ち上げ時から本修士論文の作成に至るまで，数多くのご助言を賜りました．本研究の進行において，本筋から逸れないように幾度もご指摘をいただきましたほか，研究発表の機会を多く設けてくださり，心より感謝申し上げます．日夜を問わず原稿修正や研究の相談に乗ってくださり，自身の体調等へのご配慮まで賜りましたこと，心より感謝申し上げます．また発話データの収集において，ご協力をいたきました京都大学大学院医学系研究科の木下彩栄教授と鳥畑由香里先生，意味密度のアノテーションにおいて，ご協力いただいきました京都大学総合人間学部人間・環境学研究科の永井宥之さま，岡久太郎さまここに感謝致します．最後になりましたが，2 年間お世話になりました本学ソーシャル・コンピューティング研究室の学生の皆様に深く感謝申し上げます.

参考文献

[1] 野村 総一郎，標準精神医学 第 6 版， 医学書院，2015.

[2] Snowdon, David A., et al. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the Nun Study. Jama 275.7 (1996): 528-532.

[3] 荒牧英治，若宮翔子，四方朱子木下彩栄: 話の冗長性でアルツハイマー病をみつける, 人工知能学会全国大会（JSAI）,2016.

[4] Kemper, Susan, Marilyn Thompson, and Janet Marquis. Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and semantic content. (2001).

[5] Jarrold, William, et al. Language analytics for assessing brain health: Cognitive impairment, depression and pre-symptomatic Alzheimer' s disease. Brain Informatics (2010): 299-307.

[6] Orimaye, Sylvester Olubolu, Jojo Sze-Meng Wong, and Karen Jennifer Golden. Learning predictive linguistic features for Alzheimer' s disease and related dementias using verbal utterances. Proceedings of the 1st Workshop on Computational Linguistics and Clinical Psychology (CLPsych). sn, 2014.

[7] Sirts, Kairit, Olivier Piguet, and Mark Johnson. Idea density for predicting Alzheimer's disease from transcribed speech. arXiv preprint arXiv:1706.04473 (2017).

[8] Roark, Brian, et al. Spoken language derived measures for detecting mild cognitive impairment. IEEE transactions on audio, speech, and language processing 19.7 (2011): 2081-2090.

[9] Tanaka, Hiroki, et al. Detecting Dementia Through Interactive Computer Avatars. IEEE journal of translational engineering in health and medicine 5 (2017): 1-11.

[10] Fraser, Kathleen C., Jed A. Meltzer, and Frank Rudzicz. Linguistic features identify Alzheimer' s disease in narrative speech. Journal of Alzheimer's

Disease 49.2 (2016): 407-422.

[11] Holtzman, Ronald 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 59.6 (2004): P278-P284.

[12] 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-rea gans-speech-early-sign-of-alzheimers-2015-4.

[13] 池田学, 認知症. 高次脳機能研究 (旧 失語症研究), 29(2): p. 222-228, 2009.

[14] Saczynski, Jane S., et al. The effect of social engagement on incident dementia: the Honolulu-Asia Aging Study."American Journal of Epidemiology 163.5 (2006): 433-440.

[15] Pennebaker, James W., et al. The development and psychometric properties of LIWC2015. 2015.

[16] Chand, Vineeta, et al. A rubric for extracting idea density from oral language samples. Current protocols in neuroscience (2012): 10-5.

[17] Aramaki, Eiji, et al. Vocabulary size in speech may be an early indicator of cognitive impairment. PloS one 11.5 (2016): e0155195.

[18] Brown, Cati, et al. Automatic measurement of propositional idea density from part-of-speech tagging. Behavior research methods 40.2 (2008): 540-545.

[19] The Avon Longitudinal Study of Parents and Children (ALSPAC)–study design and collaborative opportunities. European Journal of Endocrinology 151.Suppl 3 (2004): U119-U123.

[20] Ben-Zeev, Aaron. The nature of emotions. Philosophical Studies 52.3 (1987): 393-409.

[21] Li, Junyi Jessy and Nenkova, Ani. Fast and Accurate Prediction of Sentence Specificity. Proceedings of the Twenty-Ninth Conference on Artificial Intelligence (AAAI) (2015): 2281–2287.

[22] Sebastian Raschka, Python 機械学習プログラミング 達人データサイエンティストによる理論と実践, インプレス, 2016.

[23] FORBES, Katrina E.; VENNERI, Annalena; SHANKS, Michael F. Distinct patterns of spontaneous speech deterioration: an early predictor of Alzheimer's disease. Brain and Cognition, 2002, 48.2-3: 356-361.

[24] KINTSCH, Walter; KEENAN, Janice. Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive psychology, 1973, 5.3: 257-274.

[25] 飯干紀代子.【認知症の人とのコミュニケーション技法】重症度別のコミュニケーション (解説/特集). 日本認知症ケア学会誌 ,2015;14(2) :439-443.

[26] 町田明子, 上田泰士, 渥美圭大, 白土真紀, 高田定樹, 八木透. 化粧療法が高齢者の脳波にもたらす変化 認知症重症度による違い. 老年精神医学雑誌,2013;24(9):915-927.

発表リスト

- 国際会議（査読付き）

[1] Daisaku Shibata, Shoko Wakamiya, Kaoru Ito, Mai Miyabe, Ayae Kinoshita, Eiji Aramaki. VocabChecker: Measuring Language Abilities for Detecting Early Stage Dementia. Annual meeting of the intelligent interfaces community (IUI), 2018 (Demo).
[2] Daisaku Shibata, Shoko Wakamiya, Ayae Kinoshita, Eiji Aramaki. Detecting Alzheimer's Disease based on Word Category Frequencies. In Proc. of the International Conference on Computational Linguistics (COLING) Workshop on Clinical Natural Language Processing (ClinicalNLP). pp. 78-85, 2016. Osaka, Japan, December 11, 2016.

- 国内論文誌

[3] 柴田 大作，若宮 翔子，木下 彩栄，荒牧 英治. 音声発話による認知症スクリーニング技術の開発. 医療情報学 37 巻 6 号.

- 国内会議

[4] 岡久 太郎，柴田 大作，伊藤 薫，荒牧 英治，聞き手の推定した話者の語彙量と TTR との乖離に関する検討: クラウドソーシングを用いた調査，言語処理学会 4 年次大会，2018 年 3 月. クラウドソーシングを用いた調査
[5] 柴田 大作, 伊藤 薫, 若宮 翔子, 木下 彩栄, 荒牧 英治. 日本語における Idea Density: 認知症の早期発見を目指して. 第 37 回医療情報学連合大会, 2017 月 11 月 21 日.
[6] 柴田 大作, 伊藤 薫, 若宮 翔子, 宮部 真衣, 木下 彩栄, 荒牧 英治. 自由発話による認知症スクリーニングを支援するアプリケーションの開発. 第 37 回医療情報学連合大会, 2017 月 11 月 21 日.
[7] 柴田 大作，若宮 翔子，木下 彩栄，荒牧 英治. 音声発話による認知症スクリーニング技術の開発. 第 21 回春季医療情報学会春季学術大会，2017 年 6 月.

[8] 柴田 大作，若宮 翔子，伊藤 薫， 荒牧 英治. JIWC: クラウドソーシングによる日本語感情表現辞書の構築. 言語処理学会 第 23 回年次大会，2017 年 3 月.
[9] 柴田 大作，若宮 翔子，宮部 真衣，大西 正輝，山下 倫央，野田 五十樹，荒牧 英治.Twitter による群衆密度の推定－第 29 回関門海峡花火大会での実証実験－. 第 9 回データ工学と情報マネジメントに関するフォーラム, 2017 年 3 月.
[10] 柴田 大作，若宮 翔子，木下 彩栄，荒牧 英治. 単語カテゴリを用いたアルツハイマー症スクリーニングシステム. 第 36 回医療情報学連合大会，4-F-2-2，2016 月 11 月.
[11] 柴田 大作，若宮 翔子，木下 彩栄，荒牧 英治： アルツハイマーの発症に伴う代名詞の増加. 第 51 回ユビキタスコンピューティングシステム・第 5 回高齢社会デザイン合同研究発表会. Vol. 2016-UBI-51 No. 13, Vol. 2016-ASD-5 No. 13, pp. 1-7, 2016 年 8 月.