一般口演
一般口演5
知識工学
2017年11月21日(火) 08:30 〜 10:00 G会場 (10F 会議室1006-1007)

[2-G-1-OP5-2] 日本語における Idea Density: 認知症の早期発見を目指して

柴田 大作[1], 伊藤 薫[1], 若宮 翔子[1], 木下 彩栄[2], 荒牧 英治[1] （1.奈良先端科学技術大学院大学, 2.京都大学）

【背景・目的】
認知症高齢者の増加に伴い，認知症のスクリーニング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．言語能力に着目した認知症スクリーニングに関する研究は， Snowdonらの研究に始まった． Snowdonらは，ナン・スタディと呼ばれる修道女の日記を分析し，若年期の Idea Density（人文学的に定義される文章中の命題数を単語数で除した値）と晩年期における認知症の発症率には一定の関係性があることを示した．このように Idea Densityは認知症のスクリーニングに有効な指標であるにもかかわらず，日本語において検討はなされていない。そのため我々は日本語における Idea Densityの定義を行い，これ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．
【材料】
認知症と診断された患者42名（男性:7名，女性: 35名）に，最近楽しかったこと，昔楽しかったこと，好きな食べ物に関する質問を行い，その回答を録音し，書き起こしたものをコーパスとして用いる．
【手法】
日本語コーパスから Idea Densityを人手により算出することはコストが大きく，現実的ではない．そこで CPIDR (英語の Idea Densityを計算するソフトウェア)を用いる． Google社が提供している Google翻訳により英語への翻訳を行ったコーパスを CPIDR に適用することで Idea Densityの算出を行う．
【結果・考察】
算出した Idea Densityと Mini Mental State Examination (MMSE)スコアの相関係数を求めたところ，0.28と弱い相関が見られた．また，85歳以下に限定した場合，相関係数は0.57であり，中程度の相関が見られた．このことから85歳以下において，本手法は有効であり，認知症のスクリーニングに大きく寄与できる可能性が示された．

日本語における Idea Density
－ 認知症の早期発見を目指して －

柴田大作[*1]、伊藤薫[*1]、若宮翔子[*1]、木下彩栄[*2]、荒牧英治[*1]
*1 奈良先端科学技術大学院大学、*2 京都大学

Idea Density in Japanese
- Toward Early Detection for Dementia -

Daisaku SHIBATA[*1], Kaoru ITO[*1], Shoko WAKAMIYA[*1], Ayae KINOSHITA[*2], Eiji ARAMAKI[*1]
*1 Nara Institute of Science and Technology, *2 Kyoto University

[Background] The Idea Density (ID), a language based index, has been established for the detection of dementia. However, it has not been applied to non-English language because of the difficulty to convert grammatical concepts from English to another language. Instead of converting the rules of ID definition, we translate the target text data by using machine translation, which has rapidly progressed recent years. [Objective] By using automatically translated text from Japanese to English, we estimate ID. Based on indexes calculated from Japanese text, we establish the method to predict the estimated ID. [Materials] The participants of this study consisted of 42 dementia patients of 69–98 years old (mean age: 84.95). They were divided into four groups based on their age (65-74, 75-84, 85-94, 95-104). [Methods] A narratives of a patient was translated into English by using Google translation. The ID was then calculated by English ID scoring tool, Computerized Propositional Idea Density Rater (CPIDR). [Result] We confirm a significant coefficient (r = 0.568) between Idea Density and MMSE under 85 age participants (p = 0.05). We also find significant correlation (r = -0.768) between ID and the combination of multiple language ability scores (TTR + TCR + Verb ratio). [Discussion] The results have demonstrated the basic feasibility of machine translation based ID estimation. We believe that the basic concept of this translation approach could be applied the other non-English languages.

Keywords: Dementia, Idea Density, Natural Language Processing

1. はじめに

認知症者の増加に伴い、認知症のスクリーニングに関する研究が盛んに行われている。中でも自然言語処理の分野においては、認知症者の言語能力に注目し、テキストマイニングや機械学習などを用いて、認知症者と健常者を区別する研究[1)2)3)]が多く報告されている。

言語能力に注目した研究は Snowdon らの研究[4)]に始まった。Snowdonらはナン・サンディと呼ばれる修道女の日記を分析し、若年期における意味密度（自然科学における命題数を単語数で除した指標）と晩年期における認知症の発症リスクの間には大きな関係があることを示した。

意味密度を自動的に計算する研究も多く報告されている。Kemperらが開発したCPIDR[5)]は37個のルールに基づき意味密度を計算するソフトウェアである。CPIDR はいくつかの研究で利用[6)7)]されており、その信頼性は高い。

このように意味密度に関する研究は海外で盛んであるが、日本語では報告されていない。これには以下のような理由が考えられる。

1. 英語では意味密度の明確な計算方法が定義されているが、日本語ではなされていないため
2. CPIDR などの意味密度を計算する手法は英語以外の言語に対応していないため
3. 日本語と英語では文法構造が大きく異なり、英語で確立されている手法を日本語に適用することは困難であるため。

そこで我々は、日本語における意味密度の推定方法を提案する。具体的には、認知症者の日本語コーパスを機械翻訳により英語に翻訳し、CPIDR を適用することで意味密度を計算する。そして日本語において確立されている言語指標との相関関係を調査することで、日本語における意味密度の推定を実現する。

2. 実験材料

本研究では認知症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

参加者の詳細なデータを表 1 に示す。

表 1 自由会話コーパスの詳細

	中度認知症	軽度認知症
	認知症と診断され，MMSEの結果が21点以	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

研究参加者のリクルートは、研究協力者の京都大学大学院医学研究科にて行った。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[1]を経て実施された。

3. 意味密度

意味密度は Kintsch らにより提唱された言語指標で、工学的に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、その例を図 1 に示す。

Example:
The old gray mare has a big nose.

Propositions:
1. Mare is old.
2. Mare is gray.
3. Mare has nose.
4. Nose is big.

4 propositions ÷ 8 words = 0.500 idea density

図 1 意味密度の定義 [5)]

命題数の測定には、専門家によるアノテーションが必要であり、コストが大きい。そのため、英語のテキストにおいて、意味密度の計算を自動で行うソフトウェア CPIDR(Computerized Propositional Idea Density Rater) [5)]が公開されており、いくつかの研究 [6)7)]で使用されている。そのため本研究でもこれを用いて、意味密度を計算する。

4. 方法

日本語で確立されている言語指標と意味密度の相関を調査する。日本語における意味密度の計算と推定方法は以下の手順(図 2)に沿って行う。

[1] 承認番号: C1152

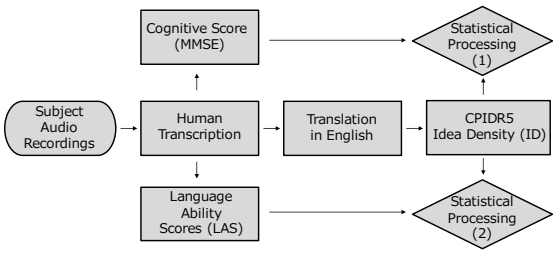

図 2 日本語における意味密度の推定方法

Step1: コーパスの翻訳

意味密度を計算するソフトウェアである CPIDR は英語のみに対応しているため、コーパスの英語翻訳を行う。翻訳には表 2 に示す 3 種類の翻訳サービスを利用する。複数の翻訳サービスを利用する理由はサービスごとに翻訳手法が異なるためである。

表 2 使用する翻訳サービス

Type	Company	Methods
Google Translation	Google	Neural Machine Translation
Bing Translation	MicroSoft	Statistical Machine Translation
excite Translation	excite	Rule Base Machine Translation

Step2: CPIDR への適用

意味密度の計算方法はいくつか報告 [5) 8)]されているが、本研究では CPIDR を利用する。

Step1 で英語への翻訳を行なったコーパスを CPIDR に適用することで意味密度を算出する。CPIDR を文章に適用する例を表 3 に示す。CPIDR はテキストから品詞タグの数を計算し、37 個のルールを適用することで、意味密度を計算するソフトウェアである。

表 3 CPIDR への適用例

Idea Density	English	Japanese
0.38	I like apples.	私はりんごが好きです.
0.56	The mare has a big nose.	鼻の大きい馬.
0.45	I watched TV yesterday.	昨日はTVを見た.

Step3: 言語指標との比較

Step2 で算出した意味密度と言語指標との相関係数を計算する。言語指標については 5 章で詳しく述べる。

5. 言語指標

本研究では以下の言語指標を用いる。

1. Type-Token Ratio (TTR)

発話した単語の延べ数(Type)と発話した単語数(Token)の比率を集計する。この値が大きいほど語彙量が多いと考えられる。単語の区切りには日本語形態素解析器(Juman++)を用いた。

$$TTR = \frac{Type}{Token}$$

2. Potential Vocabulary Size (PVS)

潜在語彙量はある人物が使用する可能性のある語彙量の予測であり、Token と Type から算出される。認知症が進行するにつれ、潜在語彙量が減少することが報告 [12)]されている。

3. Japanese Education Level (JEL)

語彙の難易度を示す指標であり、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 [9]。単語ごとに算出し、全単語の平均を JEL スコアとする。

$$JEL = \frac{Total\ of\ JEL\ value}{Number\ of\ Words}$$

4. Frequency per User popularity (FPU)

語の特殊性を示す指標である。語の特殊性は、語の出現頻度を語のユーザ数で割った値と定義する。この値が低いほど、その語は一般的であり、高いほど、その語のユーザ数が少ない語(特殊な語)であることを示す。例えば、スラングや専門用語などは高い値を持つ。 ソーシャルメディア上の 10 万人の発言を 8 ヶ月間調 査して得たデータをもとに、語のユーザ数および出現 頻度を算出し、各語の頻度・使用者数比コーパスを構築した [10]。このコーパスを参照して語ごとに頻度・ 使用者数比を算出し、全単語の値を平均した値を FPU スコアとする。

$$FPU = \frac{Total\ of\ FPU\ value}{Number\ of\ Words}$$

5. Text Compressibility Raito (TCR)

テキストを圧縮した際の圧縮率を計算した。圧縮には tar.gz 圧縮方式を適用した。

$$TCR = \frac{File\ size\ of\ Compressed\ Text}{File\ size\ of\ Original\ Text}$$

6. Emotion Ratio (ER)

発話における感情表現の使用割合を示す。感情表現は日本語感情表現辞書(JIWC) [11]を用いて算出を行う。JIWC は 7 つの感情(悲しい、不安、怒り、嫌悪感、信頼感、驚き、楽しい)にそれぞれの感情を表す感情表現が収録されている辞書である。JIWC の収録語数を表 4 に示す。

表 4 JIWC における各感情の単語数

Emotions	Number of Words
Sad	84
Anxiety	112
Angry	121
Dislike	129
Trust	97
Surprise	98
Happy	86

7. Part of Speech Ratio (PR)

発話における名詞、動詞、副詞、形容詞の使用割合を示す。品詞タグは TOKEN と同様に日本語形態素解析器(juman++)を用いて抽出を行う。

8. Kanji Ratio (KR)

発話における漢字の使用割合を示す。漢字数を全文字数で正規化することで算出を行う。具体的な計算方法を図 3 に示す。

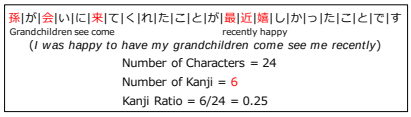

図 3 漢字割合の算出方法

6. 実験

6.1 統計処理(1):意味密度と MMSE の比較

3 種類の翻訳サービス(表 2)を用いて、サービスごとに英語コーパスを作成し、それぞれのコーパスごとで意味密度を算出した。各コーパスにおける意味密度と MMSE の関係をまとめたものを図 4、5、6 に、意味密度と MMSE の相関係数を表 5 に示す。

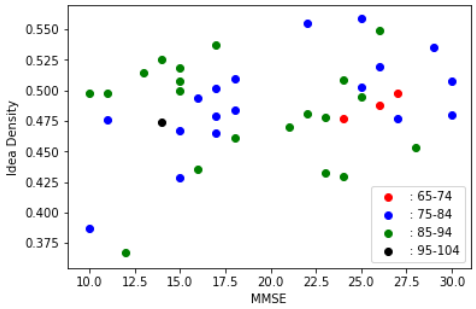

図 4 意味密度(Google 翻訳)と MMSE の関係

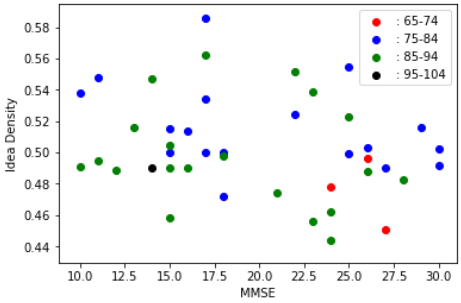

図 5 意味密度(Bing 翻訳)と MMSE の関係

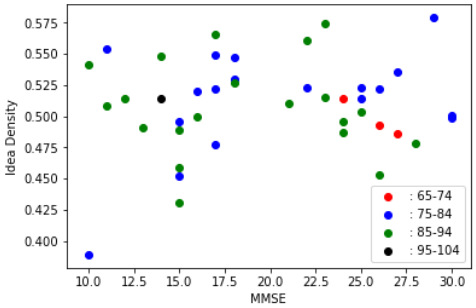

図 6 意味密度(excite 翻訳)と MMSE の関係

表 5 意味密度と MMSE の相関係数

Method	Correlation Coefficient
Google Translation	**0.279**
Bing Translation	-0.220
Excite Translation	0.125

表 5 から、Google 翻訳を用いてコーパスを翻訳し、意味密度を算出したデータが MMSE との相関係数が最も高いことがわかった。そのため本研究では Google 翻訳から算出した意味密度を用いる。

6.2 年齢を考慮した分析

一般的に認知症の有病率は 80-84 歳では 20%であるが、85

歳以上では 40%に倍増することが報告 [13]されている。そのため、早期スクリーニングという観点においては 84 歳以下に焦点を当てることも重要である。そのため、コーパスを 65-84 歳（n=21）、65-94 歳（n=33）、65-104 歳（n=42）に分け、相関係数を算出したものを表 6 に示す。

表 6 年齢を考慮した意味密度と MMSE の相関係数

Age	Correlation Coefficient
65-84	**0.568**
65-94	0.276
65-104	0.279

6.3 統計処理（2）：言語指標による意味密度の推定

意味密度と 6 章で述べた言語指標の相関係数を表 7 に示す。

表 7 意味密度と言語指標の相関係数
（相関係数が 0.400 より大きいものを太字で示す。）

Language Index	Correlation Coefficient	Language Index	Correlation Coefficient
TOKEN	**0.402**	ER (happy)	0.047
TYPE	**0.477**	ER (Trust)	-0.070
TTR	**-0.440**	ER (Sad)	0.375
PAVS	-0.200	ER (Surprise)	**0.530**
JEL	0.216	PR (Noun)	**0.564**
FPU	-0.390	PR (Verb)	0.368
TCR	**-0.680**	PR (Adjective)	**-0.470**
ER (Angry)	0.141	PR (Adverb)	0.239
ER (Anx)	0.354	PR (Conjection)	**-0.430**
ER (Dislike)	0.246	KR	0.319

加えて、意味密度と言語指標の線形和（3 種類の組み合わせ）との相関係数を算出した。相関係数が大きかった上位 5 つの組み合わせを表 8 に示す。

表 8 意味密度と言語指標の線形和の相関係数（上位 5 つ）

Index1	Index2	Index3	Correlation Coefficient
TTR	TCR	PR (Verb)	-0.786
TTR	TCR	PR (Adjective)	-0.773
TTR	TCR	ER (Angry)	-0.771
TTR	TCR	PR (Conjection)	-0.766
TTR	TCR	ER (Anx)	-0.766

7. 考察

発話時間は最大で 3 分としており、TOKEN は流暢性とみなすことができ、滑らかの話せる高齢者ほど複雑な文章を作れる可能性がある。TYPE に関しても同様に、TYPE は TOKEN が増加すると、ある程度はそれに従って増加するので、流暢性と考えることができる。

TCR は冗長性を示す指標であるが、負の相関となっている。自由発話において、次に話すことが出てこない時に同じ内容を繰り返す可能性があるので、こちらも一種の流暢性とみなすことができると考えられる。

CPIDR では動詞、形容詞、接続詞の数を利用して意味密度を計算しており、名詞は利用していない。しかし名詞は意味密度と正の相関、形容詞と接続詞は負の相関となっている。これは機械翻訳を用いたことが原因の 1 つではないかと思われる。

TTR と Surprise は現時点で考察ができておらず、今後の課題としたい。

8. まとめ

日本語における意味密度を推定するために、複数の言語指標を組み合わせることによる推定手法を提案した。Google 翻訳を用いてコーパスを翻訳し、CPIDR により意味密度を算出した場合、意味密度と MMSE の相関係数は 0.279 であった。また、85 歳以下に限定した場合、相関係数は 0.568 となった。そして、意味密度と最も相関する言語指標は TTR、TCR、PR の組み合わせであり、相関係数は-0.786 と強い負の相関が確認された。従来、日本語において意味密度を計算することは困難であったが、我々の手法を用いることで簡易的な推定が可能であると考えられる。

謝辞

本論文の作成にあたり、研究参加者のリクルート、データの収集などの作業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。

本研究の一部は JST CREST 課題番号：JPMJCR16E2、JSPS 科研費 JP16H06395、JP16H06399、JP16K12489 の支援を受けたものです。

参考文献

1) Fraser, K. C., Meltzer, J. A. and Rudzicz, F. Linguistic Features Identify Alzheimer's Disease in Narrative Speech. Journal of Alzheimer's Disease 2015; Vol. 49, No. 2, 407–422.
2) Orimaye, S. O., Wong, J. S.-M. and Golden, K. J. 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) 2014; 78–87.
3) Maria Yancheva, Frank Rudzicz. Vector space topic models for detecting Alzheimer's disease. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics 2106; 2337–2346.
4) Snowdon D.A, Kemper SJ, Mortimer JA, et al. Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. Findings from the Nun Study. JAMA 1996; 275(7): 528-532.
5) Cati Brown, Tony Snodgrass, Susan J. Kemper, Ruth Herman, Michael A. Covington. Automatic measurement of propositional idea density from part-of-speech tagging. 2008; Behavior research methods 40(2): 540–545.
6) Brian Roark, Margaret Mitchell, John Paul Hosom, Kristy Hollingshead, Jeffrey Kaye. Spoken Language Derived Measures for Detecting Mild Cognitive Impairment. 2011; 19(7) IEEE transactions on audio, speech, and language processing: 2081–2090.
7) William L. Jarrold, Bart Peintner, Eric Yeh, Ruth Krasnow, Harold S. Javitz, Gary E. Swan. Language analytics for assessing brain health: Cognitive impairment, Depression and pre-symptomatic Alzheimer's disease. 2010; In Proceedings of the 2010 In- ternational Conference on Brain Informatics: 299–307.
8) Kairit Sirts, Oliver Piguet, Mark Johnson. Idea density for predicting Alzheimer's disease from transcribed speech. 2017; In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics. 322–332.
9) 砂川有里子:学習辞書編集支援データベース作成について『学習辞書科研』プロジェクトの紹介. 2012; Vol.24 日本語教育連 絡会議論文集: 164−169.
10) 荒牧英治，増川佐知子，森田瑞樹，保田祥．日本人のオンライン・コミュニケーション上での平均使用語彙数は 8,000 語である．

2102; 情報処理学会 第 208 回 自然言語処理研究会.
11) 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. 2017; 第 21 回春季医療情報学会春季学術大会:76-77.
12) 四方朱子, 荒牧英治. 言語能力検査としての言語処理:長期間のブログ執筆を継続した認知症の 1 例. 2014; 言語処理学会 第 20 回年次大会.
13) 朝田隆. 厚生労働科学研究費補助金 認知症対策総合研 究事業 認知症の実態把握に向けた総合的研究. 2011; 平成 21 年度～平成 22 年度総合研究報告書.