

認知症者の発言は圧縮すると小さなサイズになる:

検査時の会話の圧縮率を用いたアルツハイマー病スクリーニング

Dementia Narrative becomes Compact:
Alzheimer Disease Screening using Narrative Compression Ratio in MMSE Tests荒牧英治^{*1}
Eiji ARAMAKI若宮翔子^{*1}
Shoko WAKAMIYA四方朱子^{*1}
Shuko SHIKATA木下彩栄^{*2}
Ayae KINOSHITA^{*1} 奈良先端科学技術大学院大学 ^{*2} 京都大学医学系研究科
Nara Institute of Science and Technology #1 Graduate School of medicine Kyoto University #2

Several methods have been established for the early detection of dementia, such as blood testing. However, such methods are often quite invasive both physically and mentally. In such a situation, language ability based screening techniques are awaited. This study challenges to quantify the speech of Alzheimer's disease patients and to examine the relationship between Alzheimer disease and language ability by using text compression technique. We have examined the language abilities of the speeches of healthy elders and Alzheimer's disease (derived by MMSE scores). The text compression ratio and type token ratio showed the relationship with Alzheimer's disease. It may indicate the possibility of screening.

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¹とも算定されており、国家的問題となっている。現在、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ついて調査する研究が注目されている。

今日に至るまで、認知症における失語に関する研究自体は数多くなされてきた。コミュニケーション能力 (Bayles, Tomoeda, and Trosset 1992), 文理解 (Kempler et al. 1998), 失語 (Ecklin and Schonenberger 1992; Nicholas et al. 1985), 話し言葉の変化 (Benke et al. 1990), カテゴリ化 (Nebes and Halligan 1999; Price and Grossman 2005), 単語想起 (Salmon, Heindel, and Lange 1999), 文の追唱 (Small, Kemper, and Lyons 1997) などと認知症との関連が調べられてきた。さらに、近年では、診断の補助的手段としての可能性も検討されはじめている (Tsantali, Economidis, and Tsolaki 2013)。

なかでも、Snowdon らは、発言にどれくらいの情報量が詰め込まれているかを表す尺度として意味密度 (Snowdon et al.

1996) を提案し、一定語あたりに含まれる命題数によって、認知症患者をスクリーニングしている。ただし、この命題の認定は、言語学の専門家の判断によってなされている。また、これを簡易に自動測定するツール (CPIDR (Brown et al. 2008)) も存在するが、英語言語にしか対応していない。なお、日本語においては命題の定義さえ明確化されておらず、これをそのまま適用することはできない。

本研究では、自発的な発話から日本語能力を調べ、認知症患者をスクリーニングするために、これまで意味密度と呼ばれていた指標に代替し、テキスト圧縮率を用いた手法を提案する。Zip 圧縮をはじめとした圧縮アルゴリズムの多くは、辞書法と Huffman 符号化を組み合わせたものである。この性質上、繰り返しの文字列パターンが多く含まれていれば、それだけ圧縮率は高くなる。逆に、規則化できない情報量の高い文章の圧縮率は低くなる。本稿では、圧縮率と先行研究における意味密度 (意味的な密度) は対応していると仮定し、発言の圧縮率と認知症の程度の相関を調査する。

2. 材料

材料となる認知症者の語りとして、病院を受診中の患者が、心理検査時に話す内容を収録し、描き起こしたテキストを用いた。

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)

連絡先: 荒牧英治, 〒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 aramaki@is.naist.jp

¹ http://www.keio.ac.jp/ja/press_release/2015/osa3qr00000wfbw-att/20150529_02.pdf

【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。

会話の収録は民生機 IC レコーダー (Roland 社) を用い、これを書き起こしたものをコーパスとした。

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、患者への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」²の倫理審査の承認を経て実施された。

2.3 評価

評価は MMSE (Mini Mental State Examination) によって行う。MMSE とは、1975 年に考案された知能検査であり、アルツハイマー症などの疑いがある研究協力者のために作られた簡便な検査方法である。研究協力者に対し、11 問の口頭質問を行い、各質問にあらかじめ定められた点数に基づき 30 点満点で判定する。点数の解釈として、27~30 点のときは正常、22~26 点のときは軽度認知症の疑いあり、21 点以下の場合には認知症の疑いが強いとされる。

表 1: 実験参加者。

	AD (HDS-R score of 21 or lower)	Healthy (MMSE score of 22 or higher)
性別	Men: 1; Women: 7	Men: 4; Women: 5
年齢	77.2	76.6
MMSE Score	17.0	25.1

3. 方法

言語解析には、以下の 9 つの尺度を用いる。

● **TIME**

発話した総時間。

● **TOKEN**

発話した単語の述べ数を集計する。単語の区切りには、日本語形態素解析器 (Kurohashi et al. 1994) を用いた。

● **TYPE**

発話した単語の種類数を集計する。

● **Type Token Ratio (TTR)**

前述の Type と Token の比率 (Type/Token) 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。

$$TTR = \text{Type} / \text{Token}$$

● **Token Per Second (TPS)**

1 秒あたりの発話単語数を集計する。

● **Lexical Education Level (LEL)**

単語の難易度。難易度のスコアは、日本語学習辞書を用いて決定した。日本語学習辞書では、主な日本語 17,928 語について、初級、中級、上級との分類がなされている。このうちの名詞を対象に、中級以上の名詞の割合を算出し、スコア化した。

$$LEL = \text{中級以上の名詞数} / \text{名詞数}$$

● **Average Word User (AWU)**

平均の単語ユーザ数。語のユーザ数は文献 (Aramaki et al. 2013) の値を用いた。

$$AWU = \sum_{w \in W} WU(w)$$

ここで、W は文に含まれる単語集合、WU(w) はある単語 w のユーザ数とする。

● **Average Dependency Distance (ADD)**

平均の係り受け距離を計算した。

$$ADD = \sum_{p \in P} DD(p)$$

ここで、P は文に含まれる文節の集合、DD(p) はあるフレーズ p とかかり先との間のフレーズ数 (ギャップ) とする。図 1 の例では、5 つの係り受けがあり、そのうちの 1 つにギャップがある。よって、ADD=0.2 (1/5) となる。

● **Text Compressibility Ratio (TCR)**

テキストを圧縮した際の圧縮率を計算した。圧縮には、表 2 に示す 4 種類の圧縮方式を適用した。

$$TCR = \text{圧縮後のファイルサイズ} / \text{圧縮前のファイルサイズ}$$



図 1: 係り受け距離。

表 2: 4 種類の圧縮方式。

圧縮方式	説明
ZIP	現在使われている圧縮方式である。自己解凍形式で解凍ソフトが必要ない。
LZH	フリーソフトウェア「LHA」で使用されていた形式で、主に国内で普及している。Windows 用の圧縮形式としては、Zip 形式と並んで広く利用されている。圧縮率に優れている。
TARGZ	Linux でよく使われるファイル圧縮方式である。
CAB	Microsoft 社が開発したファイル圧縮方式である。他の圧縮形式よりも圧縮率が全体的に高いといわれている。

² <http://www.ec.med.kyoto-u.ac.jp>

4. 実験

認知症群と対象者群の発話から得られた各尺度の値を比較した。その結果を表 3 に示す。発話時間(TIME)、語の数(TOKEN)や種類(TYPE)のいずれにおいても、疾患との関連は見られなかった。しかし、発話数と種類の比(TTR)においては、患者群で有意に少なくなっており、疾患が語彙の少なさと関連することを示している。

また、圧縮率(TCR)に関しても、4つある圧縮方式のうちの2つ(LZH と CAB)が相関を見せており、認知症群の発話はより小さく圧縮できることを示している。特に、LZH に関しては最も大きな有意差を示しており、現状の言語指標では最も弁別能が高い。圧縮の方式と症状との関連性に関しては、今後さらなる調査が必要である。

表 3: 実験結果.

尺度	AD	nonAD	p value
TIME	17.40	13.27	0.23
TOKEN	<u>1225</u> (518)	<u>814</u> (123)	0.08
TYPE	283.3 (105.4)	241.5 (78.0)	0.35
TTR	<u>0.249</u> (0.054)	<u>0.313</u> (0.052)	0.02
TPS	1.91 (0.99)	2.68 (0.37)	0.33
LEL	2.171 (0.157)	2.271 (0.135)	0.16
ADD	0.661 (0.20)	0.704 (0.182)	0.64
AWU	69736.1 (5877.8)	62430.4 (5578.2)	0.33
TCR _{ZIP}	<u>0.348</u> (0.054)	<u>0.395</u> (0.044)	0.054
TCR _{LZH}	<u>0.355</u> (0.04)	<u>0.402</u> (0.032)	0.015
TCR _{TARGZ}	0.389 (0.066)	0.446 (0.051)	0.060
TCR _{CAB}	<u>0.364</u> (0.048)	<u>0.413</u> (0.038)	0.029

5. 考察

本研究により、認知症者の語りを圧縮すると、健常者よりも小さくなることが示された。先行研究においては、命題密度という言語学由来の尺度が導入されていたが、もし圧縮率がこれに相当するのであれば、極めて簡易に測定することが可能となり、臨床への導入が容易となるであろう。

また、従来から、認知症者の語りについては、同じことばかり繰り返して言う、「あれ」や「それ」などの指示詞が多用されるなど、多くの言語的な特徴が示唆されてきた。圧縮率は、直感的に、これらの定性的な特徴を定量化できる可能性がある。

さらに、同様に弁別能がある発話数と種類の比(TTR)については、サンプル・サイズの影響を受けやすく、サンプル・サイズが小さい場合には大きな値になりやすい(Maas 1972) (Richards 1987))。同様に、圧縮率についても、発話サイズが小さい場合には、辞書化が十分に行われなため、同様の問題が生じている可能性があり、今後、さらなる調査が必要である。実際に、本実験でも、認知症者群の方の発話時間(TIME)や発話数

(TOKEN)が健常者群よりも大きく、このバイアスの影響を一定量受けていると思われる。しかし、これら(TIME や TOKEN)単独では、認知症者と健常者には有意な差が見られず、発話数と種類の比(TTR)や圧縮率(TCR)でのみ差が生じている点は、依然として、言語指標によるスクリーニングの可能性を示唆しているといえる。

6. おわりに

本研究は、自然言語処理技術を用いて、患者の語りを定量化し、認知症との関係を調査した。MMSE テストにより認知症疑いがある高齢者群とそうでない群の語りに対して、発話の圧縮率を調査した。結果、有意に両者を区別でき、認知症を一定の精度でスクリーニングできる可能性が示された。このような単純な方法で、認知症をスクリーニングすることが可能であるならば、今後の認知症早期発見にとって、重要な意義を持つといえる。今後、より大規模な調査でこれを確認する必要がある。

参考文献

- Aramaki, Eiji, Sachiko Maskawa, Mai Miyabe, Mizuki Morita, and Sachi Yasuda. 2013. "A Word in a Dictionary is used by Numerous Users." In *International Joint Conference on Natural Language Processing (IJCNLP2013)*. Japan.
- Bayles, K. A., C. K. Tomoeda, and M. W. Trosset. 1992. 'Relation of linguistic communication abilities of Alzheimer's patients to stage of disease', *Brain Lang*, 42: 454-72.
- Benke, T., B. Andree, M. Hittmair, and F. Gerstenbrand. 1990. '[Speech changes in dementia]', *Fortschr Neurol Psychiatr*, 58: 215-23.
- Brown, C., T. Snodgrass, S. J. Kemper, R. Herman, and M. A. Covington. 2008. 'Idea density from part-of-speech tagging', *Behav Res Methods*, 40: 540-5.
- Ecklin, R., and P. M. Schonenberger. 1992. '[Verbal ability of Alzheimer's disease patients and very elderly control patients using the Aachen aphasia test]', *Schweiz Arch Neurol Psychiatr*, 143: 371-9.
- Kempler, D., A. Almor, L. K. Tyler, E. S. Andersen, and M. C. MacDonald. 1998. 'Sentence comprehension deficits in Alzheimer's disease: a comparison of off-line vs. on-line sentence processing', *Brain Lang*, 64: 297-316.
- Kurohashi, Sadao, Toshihisa Nakamura, Yuji Matsumoto, and Makoto Nagao. 1994. "Improvements of Japanese Morphological Analyzer JUMAN." In *The International Workshop on Sharable Natural Language Resources*, 22-38.
- Maas, HD. 1972. 'Zusammenhang zwischen Wortschatzumfang und Länge eines Textes', *Zeitschrift für Literaturwissenschaft und Linguistik*, 8: 73-79.
- Nebes, R. D., and E. M. Halligan. 1999. 'Instantiation of semantic

-
- categories in sentence comprehension by Alzheimer patients', *J Int Neuropsychol Soc*, 5: 685-91.
- Nicholas, M., L. K. Obler, M. L. Albert, and N. Helm-Estabrooks. 1985. 'Empty speech in Alzheimer's disease and fluent aphasia', *J Speech Hear Res*, 28: 405-10.
- Price, C. C., and M. Grossman. 2005. 'Verb agreements during on-line sentence processing in Alzheimer's disease and frontotemporal dementia', *Brain Lang*, 94: 217-32.
- Richards, B. 1987. 'Type/Token Ratios: what do they really tell us?', *J Child Lang*, 14: 201-9.
- Salmon, D. P., W. C. Heindel, and K. L. Lange. 1999. 'Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: implications for the integrity of semantic memory', *J Int Neuropsychol Soc*, 5: 692-703.
- Small, J. A., S. Kemper, and K. Lyons. 1997. 'Sentence comprehension in Alzheimer's disease: effects of grammatical complexity, speech rate, and repetition', *Psychol Aging*, 12: 3-11.
- Snowdon, D. A., S. J. Kemper, J. A. Mortimer, L. H. Greiner, D. R. Wekstein, and W. R. Markesbery. 1996. 'Linguistic ability in early life and cognitive function and Alzheimer's disease in late life. Findings from the Nun Study', *JAMA*, 275: 528-32.
- Tsantali, E., D. Economidis, and M. Tsolaki. 2013. 'Could language deficits really differentiate Mild Cognitive Impairment (MCI) from mild Alzheimer's disease?', *Arch Gerontol Geriatr*, 57: 263-70.