Detecting Dementia from Repetition in Conversational Data of Regular Monitoring Service

Kaoru Shinkawa[1], **Keita Shimmei**[1,2], **Yasunori Yamada**[1]
[1]IBM Research - Tokyo, Tokyo, Japan
19-21 Nihonbashi Hakozaki-cho Chuo-ku, Tokyo, Japan
[2]Department of Preventive Medicine and Public Health, Graduate School of Medicine, Keio University
35 Shinanomachi, Shinjuku-ku, Tokyo, Japan
kaoruma@jp.ibm.com, shinmeikeita@keio.jp, ysnr@jp.ibm.com

Abstract

Language dysfunctions are recognized as prominent signs of dementia, and previous computational studies have shown that measuring such dysfunctions can serve as a sensitive index of cognitive decline. These features of measuring language dysfunctions have been investigated in conversational data collected during neuropsychological tests but not in data collected during daily conversations. In this study, we used data obtained from a daily monitoring service for eight elderly people, including two who had been reported as having dementia, and investigated the features that characterize repetition in conversations on different days as well as single conversations on the same day. Through the analyses, we found that features for measuring repetition significantly increase for dementia patients in terms of topic and words. The results suggest that using the repetition features over the regular conversational data is a promising approach for detecting dementia sufferers.

Keywords: Monitoring Service, Linguistic Dysfunctions, Daily Conversation, Topic Similarity, Vocabulary Richness

1. Introduction

As the world's elderly population increases, the number of people living with dementia is rising rapidly, making dementia an increasingly serious health and social problem (Prince et al., 2013). Globally, around 47 million people were living with dementia in 2015, corresponding to about 7.6% of the world's over-65-year-olds (Prince et al., 2013). Although dementia is the fifth-biggest cause of death in high-income countries, it incurs the highest annual global cost to manage (estimated to be as high as US$818 billion) because patients require constant and costly care for years (Dolgin, 2016; Prince et al., 2015). Japan is one high-income country facing a severe aging problem. The prevalence of dementia for persons 65 years or older is estimated at around 15%, and the annual cost spent on care for dementia patients was around US$120 billion (14.5 trillion JPY) in 2014 (Shikimoto et al., 2016).

Technological innovations in monitoring services for older adults as well as dementia care are expected to help people with dementia and their carers. These include diagnostic, monitoring, assistive, therapeutic, and carer supporting technologies (Livingston et al., 2017). In particular, interest is growing in technologies for early diagnosis as a possible way of improving dementia care because of recent failures in both clinical trials and laboratory work (Sperling et al., 2011). However, many people with dementia remain undiagnosed, and diagnostic coverage worldwide remains low (Prince et al., 2016). Even in high-income countries, only 40-50% of dementia sufferers have received a diagnosis (Prince et al., 2016). For example, only 45% of dementia sufferers in the United States have received clinical cognitive evaluations (Kotagal et al., 2015). Timely diagnosis is a prerequisite for good dementia care and helps people benefit from interventions, social support, and treatments. From this perspective, monitoring technology able to detect early signs of dementia in everyday situations might have great potential for supporting earlier diagnosis. Available published work has shown the usefulness of monitoring technologies for inferring an individual's state, such as stress (Lu et al., 2012), and mental fatigue (Yamada and Kobayashi, 2017a; Yamada and Kobayashi, 2017b); assessing behavioral characteristics, such as sleep quality (Rahman et al., 2015) and activities (Cook, 2010); and screening for diseases, such as bipolar disorder (Faurholt-Jepsen et al., 2016) and Parkinson's diseases (Tsanas et al., 2010). However, dementia remains difficult to detect from data collected on a daily basis for various reasons, such as people misconstruing the symptoms as a normal part of ageing.

One of the most promising ways to assess the health of dementia patients in everyday situations is identifying the evolution of language change over the course of dementia's progression. Although the most typical symptom of dementia is memory impairment due to the medial temporal lobe shrinking (Kirshner, 2012; MacKay et al., 2008), dementia is characterized by a decline from a previously attained level of performance in one or more cognitive domains such as memory, learning, executive function, and language (American Psychiatric Association, 2013). As for language function, both retrospective analyses and prospective cohort studies have shown that language problems are prevalent dating from presymptomatic periods (Van Velzen and Garrard, 2008; Oulhaj et al., 2009). In addition, studies on pathologically proven Alzheimer's disease (AD) patients have shown that they exhibited significant language changes such as syntactic simplification and impairments in lexico-semantic processing at the time of diagnosis (Ahmed et al., 2012; Ahmed et al., 2013).

Previous computational studies attempted to characterize such language dysfunctions by using acoustic, prosodic, and linguistic features extracted from data gathered while participants performed neuropsy-

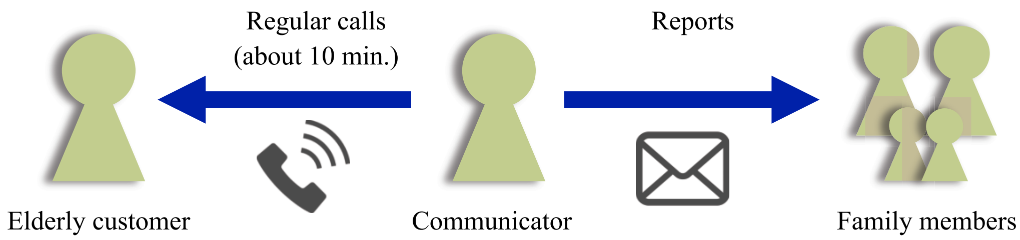

Figure 1: Overflow of regular monitoring service. A communicator calls an elderly customer once or twice a week, transcribes the conversations, and e-mails the transcripts to family members. We analyzed the conversation transcripts this service provided.

			Data duration				
Status	Gender	Age	Start	End	No. of calls	Ave. call time Mean (SD) [min.]	Ave. no. of words per conversation
Control	F	75-77	2015 Mar	2017 Apr	75	11.21 (8.85)	395.13 (124.18)
	F	80-83	2014 Jul	2017 Apr	109	16.63 (4.47)	734.34 (195.10)
	F	87-89	2016 Jan	2017 May	104	11.15 (4.46)	418.86 (235.12)
	M	66-70	2014 Jul	2017 Apr	133	10.62 (2.32)	482.89 (118.95)
	M	78-81	2014 Dec	2016 Mar	72	12.06 (2.83)	554.69 (119.03)
	M	82-85	2014 Nov	2017 Apr	226	17.75 (6.29)	572.12 (235.49)
Dementia	F	85-86	2014 Jul	2015 Nov	40	9.29 (2.15)	462.28 (204.12)
	F	88-88	2014 Jul	2014 Nov	13	7.77 (1.72)	277.94 (151.47)

Table 1: Specifications of conversational data of participants provided by the regular monitoring service.

chological tests by professionals such as medical doctors (Bucks et al., 2000; Hoffmann et al., 2010; Guinn and Habash, 2012; Fraser et al., 2016). For example, the short-term memory loss associated with dementia often brings about word-finding and word-retrieval difficulties (Henry et al., 2004; Kavé and Goral, 2018). These difficulties have typically been characterized by measuring fillers including non-words and short phrases (e.g., ”umm” or ”uh”) (Guinn and Habash, 2012; König et al., 2015). Dementia patients tend to reduce the tempo of and articulation rates in their speech (Hoffmann et al., 2010). These reductions have been measured by phonemes per second in patients' speech. Dementia patients also tend to reduce the expressiveness of their speech. This reduction has been characterized by using linguistic features such as the decrease in adjective proportion and indices related to vocabulary richness (Bucks et al., 2000; Chinaei et al., 2017). These features have been extracted from spontaneous speech data during neuropsychological tests such as image descriptions, which might be useful for characterizing everyday conversations (Tomoeda et al., 1996; Fraser et al., 2016; Chinaei et al., 2017). Studies have recently started investigating whether these language dysfunctions observed in neurodegenerative diseases including dementia can be extracted in conditions close to those of everyday life and have garnered increasing attention (Masrani et al., 2017; Shinkawa and Yamada, 2018b; Shinkawa and Yamada, 2018a). However, these studies remain limited, and further research is required to detect dementia from language data gathered from everyday situations such as social media posts, family conversations, and conversations during monitoring services for older adults.

In this study, we analyzed conversational data of older adults with and without dementia obtained from a regular monitoring service for elderly people in Japan. We focused on repetition in conversations on different days in addition to single conversations on the same day on the basis of previous observational and descriptive studies that reported atypical repetitions as one of the prominent characteristics observed in dementia patients in everyday conversations (Cook et al., 2009). The analyses revealed that features for measuring repetition of both words and topics on different days increase for people with dementia compared with controls.

2. Materials & Methods

To gain insight into how dementia affects language features extracted from daily conversations, we analyzed conversational data collected in a regular monitoring service for elderly people. For the language features, we focused on topic and word repetition in separate conversations. In this section, we first describe the conversational data we used for analysis. We next explain how to calculate features to capture topic and word repetition on different days.

2.1. Conversational data from a regular monitoring service

We used conversational data obtained from the regular monitoring service for elderly people provided by Cocolomi Co., Ltd. (http://cocolomi.net/). Their service is intended to help families living separately to catch up on the lives of their older members. A communicator calls elderly people once or twice a week and encourages them to talk

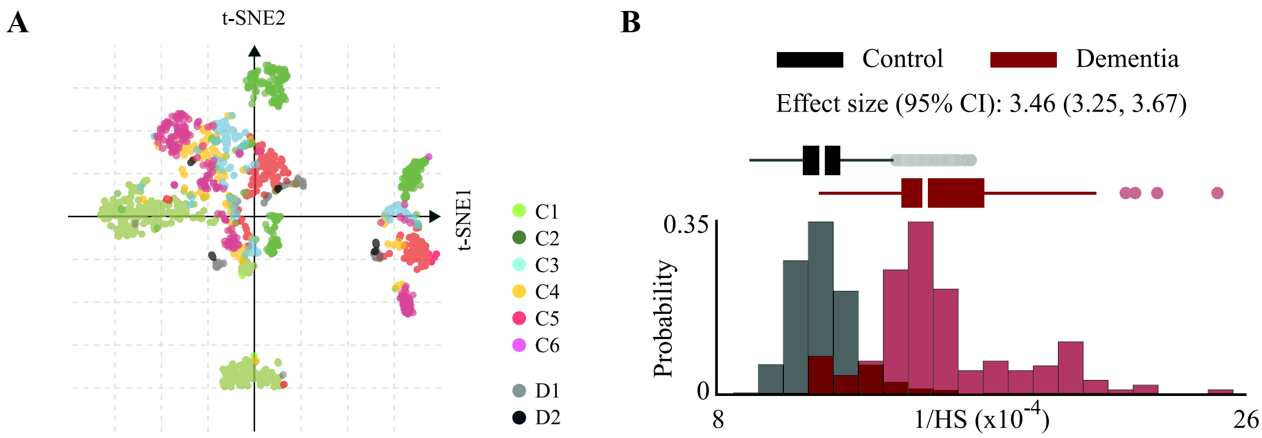

Figure 2: Representations of topic repetition. (A) Two-dimensional visualization of the topics using t-SNE. Circles are positioned on the basis of their position in the t-SNE1 and t-SNE2 dimensions. Circles represent each topic. Grey and black circles are topics extracted from the dementia participants' conversations, and the others are from controls' conversations. (B) Histogram and boxplot of the feature related to topic repetition on different days. Boxes denote the 25th (Q1) and 75th (Q3) percentiles. The line within the box denotes the 50th percentile, while whiskers denote the upper and lower adjacent values that are the most extreme values within Q3+1.5(Q3-Q1) and Q1-1.5(Q3-Q1), respectively. Filled circles show outliers.

about their daily life. The conversation is transcribed by the communicator and e-mailed to the family members (Figure 1). The communicator typically transcribes the conversation in a spoken word format omitting incomplete words and fillers.

We used the transcribed texts collected from eight Japanese people (five females and three males aged 66-89 years, i.e., 82.37 $\pm$ 5.91 years old). Two of them were reported by their families as suffering from dementia. Table 1 shows the duration the service was used the number of the reported calls, the average duration of each call, and the average word length of each report. In total, 458,738 words in 772 documents were used for the analysis. All reports were written in Japanese.

For preprocessing, we performed word segmentation, part-of-speech tagging, and word lemmatization on the transcribed texts. Required part-of-speech words were extracted from the transcribed texts, and predefined stop words were eliminated. We used the Japanese morphological analyzer MeCab (Kudo, 2005).

2.2. Feature related to topic repetition on different days

To obtain a feature related to topic repetition, we first extracted N topics from conversational data of three successive phone calls that were arranged in time sequence of data collection date during the monitoring service. We then calculated topic repetition by using topic similarities between the two sets of conversational data.

To extract topics, we used latent Dirichlet allocation (LDA), an unsupervised Bayesian probabilistic model commonly used in text analysis (Blei et al., 2003). It assumes that all documents are probabilistically generated from a set of N topics, where each topic is a multinomial distribution over the words (β) and the documents are a mixture of these topics (θ). LDA assumes every document in the corpus is generated using the following generative process:

1. A document specific topic distribution $\theta_c \sim Dir(\alpha)$ is drawn,

2. and for the ith word in the document; a topic assignment $z_i \sim \theta_c$ is drawn, and a word $w_i \sim \beta_{z_i}$ is drawn and observed.

For any given data, LDA automatically infers the latent document distribution θ_c for each document $c \in D$ and the topic distribution β_k for each of the $k = \{1, ..., N\}$ topics. The probability of the ith word in a document c is:

$$p(w_i, \theta_c) = \sum_k p(w_i|\beta_k)p(z_i = (k|\theta_c)).$$

Two sets of conversational data for calculating topic similarities were picked as two to six conversational data separated. Topic similarity was measured as cosine similarity of the word probability vectors for each topic, and the word probabilities less than 0.001 were padded with zero. Topic similarities were calculated for all possible combinations of topics extracted from different set of conversational data, and their maximum values were used as measures of topic repetition. The N was set to 5 in this study.

2.3. Features related to word repetition on same day and different days

To quantify the word repetition behavior, previous computational studies focused on sentence similarities such as calculating the cosine distance between each pair of sentences in a conversation (Fraser et al., 2016). In this study, we focused on the feature of vocabulary richness as word repetition could result in a small number of distinct words being used in a conversation (Manschreck et al., 1981). Vocabulary richness was calculated by three typical measures (Bucks et al., 2000; Honoré, 1979): type-token ratio (TTR), Brunet's index (BI), and Honoré's statistic (HS). TTR is computed by dividing the total number of words

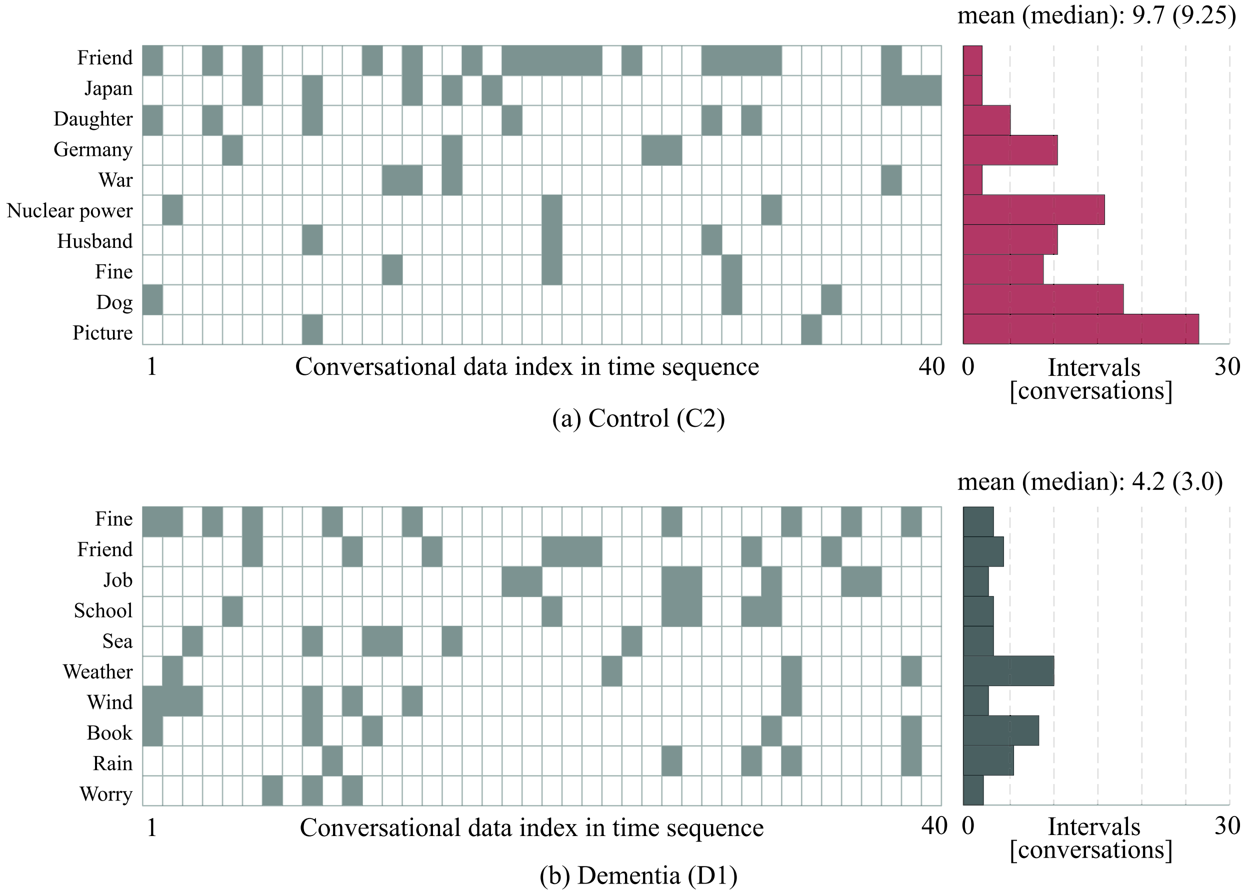

Figure 3: Appearance and interval of featured words in conversation. Featured words were translated from Japanese into English. Matrix with grey cells shows the word appearance in each conversation. The conversational data indices are ordered in time sequence from left to right. Red bars show the intervals between words.

(N) into the number of different word types (U). By using the same U and N, BI is also defined as BI=$N^{U^{-0.165}}$. Unlike with other measures related to vocabulary richness, the vocabulary richness becomes greater as BI becomes smaller. HS gives particular importance to unique vocabulary items used only once (N_{uni}). HS is defined as HS=$100 \log N/(1 - N_{\text{uni}}/U)$. Previous studies reported that HS relatively showed a higher discrimination power than other measures for detecting dementia (Fraser et al., 2016). Therefore, we focused on HS and used its inversed number as a measure of repetition.

We first obtained pairs of conversational data D_i and D_j separated by t days ($T - M \leq t \leq T + M$). D_i and D_j contain all the words except numerals and symbols. HS$_i$ and HS$_j$ were extracted from D_i and D_j by calculating HS. Next, we defined D_{ij} as a combined document of D_i and D_j and extracted HS$_{ij}$ as a feature of repetitiveness in conversations on different days. As a feature of repetitiveness in conversations on the same day, HS$_i^{-1}$ (HS$_j^{-1}$) was used.

3. Results

We investigated how topics in conversations in regular monitoring services differed between older adults with and without dementia. We applied LDA to conversational data of each participant and obtained pairs of topics and word probability vectors. We obtained 1,240 word probability vectors represented in 795 dimensions. We investigated the similarities of the word probability vectors extracted from the conversations of each individual throughout the period of the monitoring service. To this end, we used the method of t-distributed stochastic neighbor embedding (t-SNE) (Van der Maaten and Hinton, 2008), which is widely used for visualizing high-dimensional datasets. Specifically, it models each high-dimensional object by two- or three-dimensional points in such way that similar objects are modeled by nearby points and dissimilar objects are modeled by distant points. Figure 2A shows the two-dimensional representation of the word probability vectors. The topics extracted from the conversation of participants with dementia are localized near each other in each participant. This could be considered to be because similar topics appear more frequently in conversations of individuals with dementia than in conversations of those without.

We next quantified such topic similarities by using the topic-repetition feature between conversations on different days. We investigated if the feature shows the difference between individuals with and without dementia. The feature was measured by using an effect size (Cohen's d) (Nakagawa and Cuthill, 2007). For Cohen's d, a 0.8 effect size is large, 0.5 medium, and 0.2 small. We observed that the feature extracted from the conversations of individuals with dementia was significantly higher than that extracted from the conversations of individuals without (effect size of 3.46, 95% Confidence interval (CI): 3.25-3.67; Figure 2B). Topic repetition as the behavior of dementia sufferers might be captured in conversations of the daily monitoring service.

For a prior investigation of the word-repetition feature, we focused on the repetition interval of the featured words in conversations of two participants with and without dementia. We selected the top ten words on the basis of the word probabilities of topics extracted by LDA. We investigated their intervals between repetitions by calculating the mean duration of each word occurrence. From the results, the mean durations were 9.7 for the control and 4.2 for the participant with dementia (Figure 3). This result indicates that word-repetition intervals in conversations of individuals with dementia might be shorter than those in conversations of controls.

To objectively measure the repetition of words, we investigated word-repetition features in both single conversations on the same day and paired conversations on different days. We observed higher repetition in conversations of individuals with dementia than in conversations of those without in both single and paired conversational data (effect size of 1.58, 95% CI 1.28-1.88 for single conversation and effect size of 2.67, 95% CI 2.10-3.24 for paired conversation; Figure 4). The difference was larger in paired conversational data, which suggests that monitoring word repetition in conversations on different days may help to detect signs of dementia in daily life.

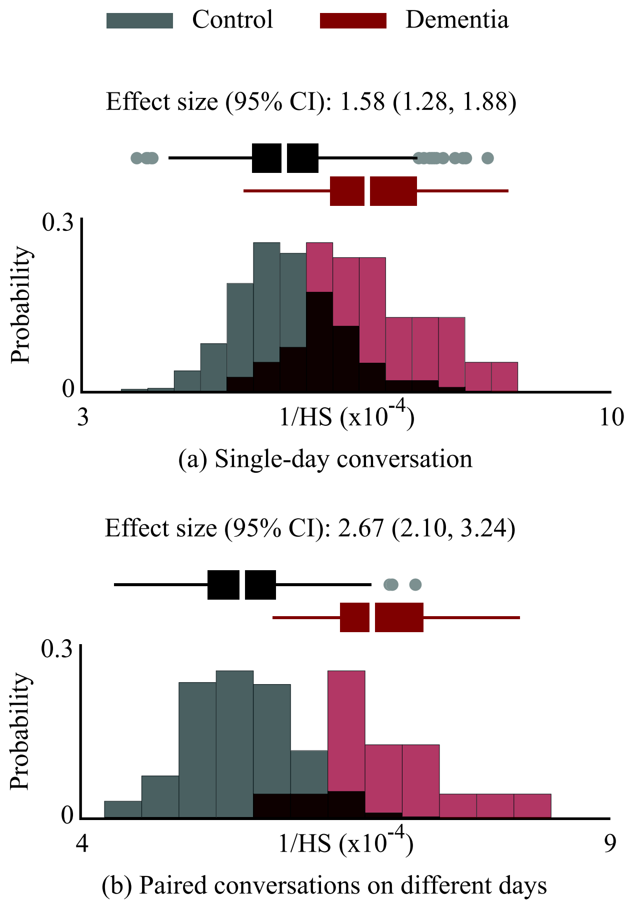

Figure 4: Comparison of histogram and boxplot between the word-repetition feature, extracted from single and paired conversations. Boxes denote the 25th (Q1) and 75th (Q3) percentiles. The line within the box denotes the 50th percentile, while whiskers denote the upper and lower adjacent values that are the most extreme values within Q3+1.5(Q3-Q1) and Q1-1.5(Q3-Q1), respectively. Filled circles show outliers.

4. Conclusion

We investigated word and topic repetition in daily conversations as a sensitive index of cognitive decline in dementia. We focused on repetition in conversations on different days in addition to single conversations on the same day on the basis of previous observational and descriptive studies that reported atypical repetitions as one of the prominent characteristics observed in dementia patients in everyday conversations.

We investigated topic- and word-repetition features by using conversational data obtained from a regular monitoring service. First, we visualized the topic probabilities using t-distributed stochastic neighbor embedding (t-SNE) to obtain an overview of topic distribution of each participant. Next, we investigated the feature of topic repetition in separate conversations on different days. We observed higher repetitiveness for participants with dementia than those without. This result indicates that the topic-repetition feature was able to capture atypical repetition in daily conversation of participants with dementia. For the word-repetition feature, we first investigated the intervals between the featured words in conversations. The intervals between repeated words were shorter in conversations of participants with dementia than in conversations of those without. Next, we investigated the word-repetition feature in single-day and different-day conversations. We observed the increase in features for measuring repetition in patients with dementia compared to those without in both topic and words. The feature related to repetition in conversations in regular monitoring services might be useful for discriminating individuals with dementia from controls.

One of the limitations in this study was its small number of participants. In future work, we will need to confirm our results on a larger number participants. Another limitation was the participants' labels for dementia sufferers and healthy controls. As mentioned in the Materials & Methods section, in this study, the participants' labels were based on not clinical assessments including dementia types and severities but reports from the participants' families.

5. Acknowledgements

We sincerely thank A. Kamiyama, Y. Masuda, J. Hayakawa, and K. Cho at Cocolomi Co., Ltd. for providing all of the data used in this study, and appreciate valuable comments and suggestions.

6. Bibliographical References

Ahmed, S., de Jager, C. A., Haigh, A.-M. F., and Garrard, P. (2012). Logopenic aphasia in Alzheimer's disease: clinical variant or clinical feature? *J Neurol Neurosurg Psychiatry*, 83(11):1056–1062.

Ahmed, S., de Jager, C. A., Haigh, A.-M., and Garrard, P. (2013). Semantic processing in connected speech at a uniformly early stage of autopsy-confirmed Alzheimer's disease. *Neuropsychology*, 27(1):79–85.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Ed. (DSM-5)*. American Psychiatric Pub.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J Mach Learn Res*, 3(Jan):993–1022.

Bucks, R. S., Singh, S., Cuerden, J. M., and Wilcock, G. K. (2000). Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology*, 14(1):71–91.

Chinaei, H., Currie, L. C., Danks, A., Lin, H., et al. (2017). Identifying and avoiding confusion in dialogue with people with Alzheimer's disease. *Computational Linguistics*, pages 377–406.

Cook, C., Fay, S., and Rockwood, K. (2009). Verbal repetition in people with mild-to-moderate Alzheimer disease: a descriptive analysis from the VISTA clinical trial. *Alzheimer Dis Assoc Disord*, 23(2):146–151.

Cook, D. J. (2010). Learning setting-generalized activity models for smart spaces. *IEEE intelligent systems*, 2010(99):1.

Dolgin, E. (2016). How to defeat dementia. *Nature*, 539(7628):156–158.

Faurholt-Jepsen, M., Busk, J., Frost, M., Vinberg, M., et al. (2016). Voice analysis as an objective state marker in bipolar disorder. *Transl Psychiatry*, 6(7):e856.

Fraser, K. C., Meltzer, J. A., and Rudzicz, F. (2016). Linguistic features identify Alzheimer's disease in narrative speech. *J Alzheimers Dis*, 49(2):407–422.

Guinn, C. I. and Habash, A. (2012). Language analysis of speakers with dementia of the Alzheimer's type. In *AAAI Fall Symposium: Artificial Intelligence for Gerontechnology*, pages 8–13.

Henry, J. D., Crawford, J. R., and Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia*, 42(9):1212–1222.

Hoffmann, I., Nemeth, D., Dye, C. D., Pákáski, M., et al. (2010). Temporal parameters of spontaneous speech in Alzheimer's disease. *Int J Speech Lang Pathol*, 12(1):29–34.

Honoré, A. (1979). Some simple measures of richness of vocabulary. *Association for literary and linguistic computing bulletin*, 7(2):172–177.

Kavé, G. and Goral, M. (2018). Word retrieval in connected speech in Alzheimer's disease: a review with meta-analyses. *Aphasiology*, 32(1):4–26.

Kirshner, H. S. (2012). Primary progressive aphasia and Alzheimer's disease: brief history, recent evidence. *Curr Neurol Neurosci Rep*, 12(6):709–714.

König, A., Satt, A., Sorin, A., Hoory, R., et al. (2015). Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimers Dement (Amst).*, 1(1):112–124.

Kotagal, V., Langa, K. M., Plassman, B. L., Fisher, G. G., et al. (2015). Factors associated with cognitive evaluations in the United States. *Neurology*, 84(1):64–71.

Kudo, T. (2005). Mecab: Yet another part-of-speech and morphological analyzer. *http://mecab. sourceforge. net/*.

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., et al. (2017). Dementia prevention, intervention, and care. *Lancet*, 390(10113):2673–2734.

Lu, H., Frauendorfer, D., Rabbi, M., Mast, M. S., et al. (2012). Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. In *Proc ACM Int Conf Ubiquitous Comput*, pages 351–360. ACM.

MacKay, D. G., James, L. E., and Hadley, C. B. (2008). Amnesic HM's performance on the language competence test: Parallel deficits in memory and sentence production. *J Clin Exp Neuropsychol*, 30(3):280–300.

Manschreck, T. C., Maher, B. A., and Ader, D. N. (1981). Formal thought disorder, the type-token ratio and disturbed voluntary motor movement in schizophrenia. *Br J Psychiatry*, 139(1):7–15.

Masrani, V., Murray, G., Field, T., and Carenini, G. (2017). Detecting dementia through retrospective analysis of routine blog posts by bloggers with dementia. *BioNLP 2017*, pages 232–237.

Nakagawa, S. and Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc*, 82(4):591–605.

Oulhaj, A., Wilcock, G. K., Smith, A. D., and de Jager, C. A. (2009). Predicting the time of conversion to MCI in the elderly role of verbal expression and learning. *Neurology*, 73(18):1436–1442.

Prince, M., Guerchet, M., and Prina, M. (2013). *The global impact of dementia 2013-2050*. Alzheimer's Disease International.

Prince, M., Wimo, A., Guerchet, M., Ali, G.-C., et al. (2015). *World Alzheimer Report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends*. Alzheimer's Disease International.

Prince, M., Comas-Herrera, A., Knapp, M., Guerchet, M., and Karagiannidou, M. (2016). World Alzheimer report 2016: improving healthcare for people living with dementia: coverage, quality and costs now and in the future.

Rahman, T., Adams, A. T., Ravichandran, R. V., Zhang, M., et al. (2015). DoppleSleep: A contactless unobtrusive sleep sensing system using short-range doppler radar. In *Proc ACM Int Conf Ubiquitous Comput*, pages 39–50.

Shikimoto, R., Sado, M., and Mimura, M. (2016). The social costs of dementia in Japan: Focusing on the informal care cost. *Brain and nerve*, 68(8):939–944.

Shinkawa, K. and Yamada, Y. (2018a). Topic repetition in conversations on different days as a sign of dementia. In *Medical Informatics Europe*.

Shinkawa, K. and Yamada, Y. (2018b). Word repetition in separate conversations for detecting dementia: A preliminary evaluation on regular monitoring service. In *Proceedings of AMIA Informatics Summit*.

Sperling, R. A., Aisen, P. S., Beckett, L. A., Bennett, D. A., et al. (2011). Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.*, 7(3):280–292.

Tomoeda, C. K., Bayles, K. A., Trosset, M. W., Azuma, T., et al. (1996). Cross-sectional analysis of Alzheimer disease effects on oral discourse in a picture description task. *Alzheimer Disease & Associated Disorders*, 10(4):204–215.

Tsanas, A., Little, M. A., McSharry, P. E., and Ramig, L. O. (2010). Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. *IEEE Trans Biomed Eng*, 57(4):884–893.

Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *J Mach Learn Res*, 9:2579–2605.

Van Velzen, M. and Garrard, P. (2008). From hindsight to insight–retrospective analysis of language written by a renowned Alzheimer's patient. *Interdiscip Sci Rev*, 33(4):278–286.

Yamada, Y. and Kobayashi, M. (2017a). Detecting mental fatigue from eye-tracking data gathered while watching video. In *Proceedings of Conference on Artificial Intelligence in Medicine in Europe*, pages 295–304. Springer.

Yamada, Y. and Kobayashi, M. (2017b). Fatigue detection model for older adults using eye-tracking data gathered while watching video: Evaluation against diverse fatiguing tasks. In *IEEE International Conference on Healthcare Informatics*, pages 275–284. IEEE.