Exploring Conversations between a Practitioner and a Person with Dementia

Kotaro Hara
kotarohara@smu.edu.sg
Singapore Management University
Singapore

Rosiana Natalie
rnatalie.2019@phdcs.smu.edu.sg
Singapore Management University
Singapore

Wei Soon Cheong
wscheong.2020@scis.smu.edu.sg
Singapore Management University
Singapore

Jingjing Gu
jjgu@smu.edu.sg
Singapore Management University
Singapore

Qianli Xu
qxu@i2r.a-star.edu.sg
Institute for Infocomm Research (I2R),
Agency for Science, Technology and
Research (A*STAR)
Singapore

ABSTRACT

In social service centers, practitioners engage in conversations with clients with dementia to facilitate their daily activities and provide support when they are distressed. However, the nature of the care demands the practitioner's active engagement, which becomes difficult to deliver as the number of people who need care expands. Researchers have been investigating the efficacy of developing agents that assume conversational tasks to alleviate this work. To contribute to the future design of agents for caregiving, we collected and analyzed ten conversations between clients with mild dementia and practitioners who provide care. Our analyses of turn-taking dynamics and dialogue acts with 15k utterances uncovered patterns such as noticeable differences in clients' and practitioners' conversational dynamics and the prevalence of neutral-toned, question-oriented utterances by practitioners. We then prototyped a large language model-based script that generates responses to client utterances. We found potential approaches and challenges for making its utterance pattern more similar to that of a practitioner.

ACM Reference Format:

Kotaro Hara, Rosiana Natalie, Wei Soon Cheong, Jingjing Gu, and Qianli Xu. 2024. Exploring Conversations between a Practitioner and a Person with Dementia. In *The 26th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '24), October 27–30, 2024, St. John's, NL, Canada.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3663548.3688523

1 INTRODUCTION

Numerous studies have explored the effectiveness of technical approaches in dementia care for older adults [11, 15–17]. Despite these advances, developing technological solutions that enable individuals with dementia to participate in meaningful activities, such as reminiscence therapy, remains a significant challenge [9]. In social service centers, practitioners employ reminiscence therapy during conversations with dementia clients (*e.g.,* to enrich their experiences) [2, 22, 30]. Although these interactions are crucial to caregiving, they require intensive practitioner engagement, which is increasingly difficult to sustain as the demand for care grows.

Previous research has investigated how autonomous agents can replicate the benefits of human-to-human interactions to improve the scalability of care for individuals with dementia [7, 8, 14, 25, 27, 28, 32]. For example, Cruz-Sandoval *et al.* created a social robot and demonstrated its positive impact on dementia-related symptoms [8]. However, the degree to which these human-agent interactions resembled human-like conversation, *e.g.,* in terms of utterance lengths and dialogue acts within caregiving scenarios has not been examined. This gap prevents us from understanding whether the previously identified positive impacts are attributed to the agent's ability to conduct human-like conversations.

As a first step to address this gap, we examined conversational patterns in reminiscence therapy between practitioners and individuals with mild dementia. We recorded ten conversations between practitioners and people with dementia and investigated how the conversations unfold in reminiscence therapy. Each conversation was approximately 30 minutes to 1 hour long. Using the transcripts of the conversations and natural language processing (NLP) techniques, we analyzed the utterance lengths and dialogue acts [29] in the conversations. Our analyses uncovered patterns such as the prevalence of certain dialogue acts by practitioners (*e.g.,* `statement-non-opinion`, `yes-no-question`). We then prototyped a large language model (LLM) based script and studied (dis)similarity of its responses to human practitioner's utterances.

2 DIALOGUE RECORDING

To collect data on conversations between practitioners and individuals with dementia, we collaborated with a local organization that provides care. We aimed to analyze conversations in the languages prevalent in the local community, English and Chinese. We recruited individuals with mild dementia as conversational care is most beneficial in the condition's early stages. We recruited practitioners from the organization to serve as conversational partners.

We formed ten conversation pairs, each comprising a client with mild dementia and a practitioner. Ten clients (8 females, 2 males; 4 English and 6 Chinese speakers) and six practitioners participated.

ASSETS '24, October 27–30, 2024, St. John's, NL, Canada

ACM ISBN 979-8-4007-0677-6/24/10.
https://doi.org/10.1145/3663548.3688523

We scheduled a recording session at a time convenient for both the client and the practitioner. We obtained informed consent from all parties. We also received assent from the clients' caregivers because the clients may have limited capacity to consent due to their cognitive condition. The recordings occurred at the organization's facility, where clients typically receive care. We conducted sessions in a dedicated, quiet area. Each session began by escorting the client to the designated space. Once the recording started, all individuals, except the practitioner and the client with dementia, were asked to leave the room to minimize distractions. The session commenced with an introductory greeting, after which the practitioner engaged the client in reminiscence therapy using various probes (*e.g.,* digital images displayed on a smartphone, physical items like historical photos of local areas, and personal artifacts provided by the clients' families). We recorded each session in video and audio, but our analysis focused on the audio recordings. To transcribe the audio recordings, we used an automated audio transcription tool [24] and used a speaker diarization tool to identify speakers and timestamp their utterances [3, 4, 23]. We then manually reviewed the transcripts to correct errors.

3 ANALYSES OF DIALOGUES AND FINDINGS

We used English-based text analysis tools for consistency and their performance [10]. This required us to translate the six Chinese transcripts into English, for which we used the publicly available `SeamlessM4T v2` model [5]. Though the automated translation was not perfect, manual inspection by a research member confirmed that most translations were generally accurate and captured the essential semantics of the text. Each processed conversation transcript contained the following data: the start and end timestamps of the transcribed text in seconds, the current speaker, and the transcribed (and translated) text. Each row of data in the transcript represents a single utterance by either the client with dementia or a practitioner. Note that both speakers could consecutively have multiple utterances, forming a single dialogue turn. For the dialogue act analysis described below, the unit of analysis was an utterance, following the past practices for dialogue act analysis [20, 29].

3.1 Utterance Length

Literature suggested the importance of clarity and brevity in practitioners' utterances in guiding dialogues [6, 21]. However, more concrete, quantitative metrics in their verbal patterns would be informative in designing future conversational systems. So, we ask: *What is each utterance's typical length and word count?*

3.1.1 Analysis Method. We performed our analysis on both utterance-level and turn-level dialogue. For the latter, consecutive utterances within each transcript were merged to form a single turn, with their start and end timings adjusted accordingly. The lengths of each level of dialogue were recorded in seconds by taking the difference between their respective end and start timings.

3.1.2 Result. Overall, practitioners contributed 629.2 utterances (SD=217.2), while clients with dementia had a higher average of 857.7 utterances (SD=401.2). That is, the clients with dementia generally had longer utterances than the practitioners. At an utterance level, clients with dementia spoke on average 2.12 seconds and 5.36 words, unlike practitioners, who averaged 1.83 seconds and 5.09 words. This trend also extended to turns; clients had an average turn length of 5.10 seconds, with 1.92 sentences and 12.3 words, while practitioners had shorter turns, averaging 3.29 seconds, 1.57 sentences, and 8.53 words. We observed some clients had extended speaking periods, in which their utterances and turns were significantly lengthier, which suggests the clients were facilitated to engage in monologues.

3.2 Dialogue Act

Dialogue acts are functional units of a conversation that represent the intent behind each utterance made by a speaker [29]. Analyzing and understanding dialogue acts in each party's utterances could help a system designer in answering: *What type of utterance should an agent emit in each conversational turn?*

3.2.1 Analysis Method. We analyzed the dialogue act of each utterance using the SWBD-DAMSL tag set, a standard for annotating shallow discourse structures [13, 29, 31]. To computationally identify utterance's dialogue act, we used `DialogTag` [19], a pre-trained machine learning model based on DistilBERT and fine-tuned on the Switchboard-1 corpus. This enabled us to assign one of the 38 dialogue act tags to each utterance. Resulting tags include, for example, `yes-no-questions` (*e.g.,* "Can I take a closer look?"), `yes-answers` (*e.g.,* "Yeah"), `appreciations` (*e.g.,* "That's great"), and others (see the complete list in [29]). To explore the distribution of dialogue acts in the utterances of both practitioners and clients, we classified the dialogue act of every utterance across the ten sessions. Subsequently, we counted the frequency of different dialogue acts of clients's and practitioners' utterances.

Dialogue Act	Client	Practitioner
Statement-non-opinion	63.2%	26.1%
Statement-opinion	6.4%	-
Collaborative completion	5.2%	-
Action-directive	2.9%	-
Yes-no-question	2.8%	17.0%
Appreciation	-	10.0%
Wh-question	-	8.0%
Acknowledge (backchannel)	-	7.3%
Total (Top 5)	80.5%	68.4%

Table 1: Comparison of five most common dialogue acts of clients and practitioners.

3.2.2 Result. Table 1 represents the five most common dialogue acts for clients and practitioners. Clients' utterances are mainly `statement-non-opinion` (63.2%), with their top five acts accounting for 80.5%. Whereas practitioners show greater diversity in their acts; while `statement-non-opinion` remains their most frequent act (26.1%), they question more (`yes-no-question`: 17.0%, `wh-question`: 8.0%) and exhibit more varied response types.

3.3 Section Summary

Analyses revealed that clients with dementia generally had longer utterances and turn lengths than practitioners. The result aligns with existing literature indicating that people with dementia tend to speak more in situations like a reminiscence therapy session [26]. Clients' longer turns, and `statement-*` tags suggested a storytelling behavior typical in reminiscence therapy. In contrast, practitioners often engaged in shorter, more active listening roles, characterized by questions, acknowledgments, and appreciations. Such a facilitative pattern aligns with what we expect from effective conversational care providers [1, 18].

4 RESPONSE GENERATING SCRIPT

To investigate the feasibility of creating an agent that imitates the behavior of a human practitioner, we used a large language model (LLM) to implement a prototype Python script that responds to client's utterance. We assessed whether such a script could generate responses that are similar to that of human practitioners.

4.1 Large Language Model-based Script

We used Python MLX library with `Phi-3-mini-4k-instruct` model to implement the script. The model was one of the most performant LLM that is small enough to train and run on a local computer. Because our informed consent promised to store the collected data only on the research team's machines, the model's relatively small size that let us run it on our local computer was a good fit.

Our script took a prompt and client's utterances to generate a response. Based on our prior knowledge and findings from the previous section, we designed the prompt for our LLM. The following prompt aimed to generate text for reminiscence therapy:

```
You are a practitioner specializing in dementia care at
a social service center. You are conducting a session
using reminiscence therapy to engage with an individual
with mild dementia. Your goal is to facilitate a
conversation that helps the person recall and discuss
pleasant memories, thereby improving their emotional
well-being. Try to use questions and gentle prompts to
encourage the individual to share their experiences and
stories from the past.
```

Furthermore, to let the model imitate the practitioner' style of utterance, we trained the model with Low-Rank Adaptation (LoRA) for fine-tuning [12], which enables domain adaptation with a small training data. We only used the four English transcripts for training and evaluation, because we did not want the machine translated text to influence the quality of text generation. We split the four English transcripts into two pairs; the first pair of transcripts were used for LoRA fine-tuning. We used the second pair for evaluation, which had 891 conversational turns where a client spoke and a practitioner responded.

4.2 Preliminary Evaluation

We compare the text generated by our LLM (with and without LoRA adaptation) to practitioners' responses to the clients' utterances. Just like the analyses from the previous section, we evaluate the volume of responses and dialogue acts of generated sentences.

4.2.1 Utterance Length. The median number of sentences per response were 1 (mean=1.4, SD=0.69) for human practitioners, 5 (5.1, 1.2) for the responses from LLM without LoRA, and 3 (2.7, 1.3) for the responses from LLM with LoRA. The median number of words in the responses were 6 (7.0, 6.1), 61 (58.7, 12.8), and 31 (29.3, 15.2), respectively. The result shows that both non-adapted and adapted LLMs over-generated text compared to human practitioners, but one without adaptation being significantly more verbose. For example, in one conversational turn, the client was talking about what they do at home; the client said, *"What I don't want to throw away, I put inside [the cabinet]."* The practitioner first acknowledged, then tried to change the topic, by responding *"Ah okay okay I see. Okay okay. So you enjoy your work a lot?"*. Unlike this succinct human response, our LLM without LoRA generated:

```
It sounds like you're sharing a memory about a specific
action or activity. Let's explore that further. Can you
tell me more about the object you mentioned? What was it,
 and what did you use it for? Sometimes, talking about
the things we used to do can help us remember and feel
good about our past experiences. Here are some prompts
to help you recall and share your memories: Do you have
any favorite childhood toys or games?
```

The text is verbose and overly explanatory. LoRA adaptation slightly suppressed the verbosity:

```
It sounds like you're trying to recall a specific
activity or event. Let's try to remember together. Do
you have any special memories of when you were doing
something similar?
```

4.2.2 Dialogue Act. To understand the conversational pattern of our LLM model, we analyzed generated sentences' dialogue acts. Like our previous analysis, we split the text into sentences and used `DialogTag` to automatically annotate the sentences. We present the frequency of dialogue acts for `statement-non-opinion`, `yes-no-question`, `appreciation`, `wh-question`, and `acknowledge (backchannel)`—the acts that we found to be most commonly used.

Table 2 shows the frequencies of the acts for the sentences generated by human practitioners, LLM without LoRA adapation, and LLM with LoRA adaptation. LLM consistently used `statement-non-opinion`, `appreciation`, and `acknowledge (backchannel)` less frequently compared to the human practitioners. On the other hand, LLM asked more `yes-no-question` than the practitioners. LLM with LoRA adaptation suppressed `wh-question` slightly. Another notable result was that LLM had more `statement-opinion` (without LoRA: 34.9%, with LoRA: 24.7%) than human (8.5%), showing that LLM are being less neutral compared to human.

4.3 Section Summary

Our script could generated text prompting clients to reminisce, but a gap persists between practitioners and our script's capabilities. The script struggled with response conciseness, though it improved after fine-tuning. The LLM (both with and without adaptation) failed to replicate practitioners' liberal use of `acknowledge (backchannel)` responses like "hm" to show attentiveness, highlighting difficulties in handling meta-conversational acts. Future work could focus on incorporating more diverse training data for fine-tuning, particularly emphasizing meta-conversational elements. As LLM-based agents approach practical viability, studying their long-term impact on real communication with clients becomes crucial.

Dialogue Act	Human Practitioner	LLM without Adaptation	LLM with Adaptation
Statement-non-opinion	24.4% (303)	17.0% (777)	12.4% (300)
Yes-No-Question	12.8% (154)	13.1% (598)	22.7% (550)
Appreciation	3.1% (39)	1.9% (85)	2.8% (67)
Wh-Question	6.1% (76)	6.2% (286)	2.4% (59)
Acknowledge (Backchannel)	22.3% (278)	0% (0)	0.4% (9)

Table 2: Frequency of dialogue acts in responses to clients.

ACKNOWLEDGMENTS

This research was supported by the SMU-A*STAR Joint Lab in Social and Human-Centered Computing | Human-AI Synergy Pillar (Grant No. SAJL-2022-HAS002/C232918002)

REFERENCES

[1] Sarah Alsawy, Sara Tai, Phil McEvoy, and Warren Mansell. 2020. 'It's nice to think somebody's listening to me instead of saying "oh shut up"'. People with dementia reflect on what makes communication good and meaningful. *Journal of Psychiatric and Mental Health Nursing* 27, 2 (April 2020), 151–161. https://doi.org/10.1111/jpm.12559

[2] Arlene J. Astell, Maggie P. Ellis, Norman Alm, Richard Dye, and Gary Gowans. 2010. Stimulating people with dementia to reminisce using personal and generic photographs. *International Journal of Computers in Healthcare* 1, 2 (2010), 177. https://doi.org/10.1504/IJCIH.2010.037461

[3] Hervé Bredin. 2023. pyannote.audio 2.1 speaker diarization pipeline: principle, benchmark, and recipe. In *Proc. INTERSPEECH 2023*.

[4] Hervé Bredin and Antoine Laurent. 2021. End-to-end speaker segmentation for overlap-aware resegmentation. In *Interspeech*. https://api.semanticscholar.org/CorpusID:233204561

[5] Seamless Communication, Loïc Barrault, Yu-An Chung, Mariano Coria Meglioli, David Dale, Ning Dong, Mark Duppenthaler, Paul-Ambroise Duquenne, Brian Ellis, Hady Elsahar, Justin Haaheim, John Hoffman, Min-Jae Hwang, Hirofumi Inaguma, Christopher Klaiber, Ilia Kulikov, Pengwei Li, Daniel Licht, Jean Maillard, Ruslan Mavlyutov, Alice Rakotoarison, Kaushik Ram Sadagopan, Abinesh Ramakrishnan, Tuan Tran, Guillaume Wenzek, Yilin Yang, Ethan Ye, Ivan Evtimov, Pierre Fernandez, Cynthia Gao, Prangthip Hansanti, Elahe Kalbassi, Amanda Kallet, Artyom Kozhevnikov, Gabriel Mejia Gonzalez, Robin San Roman, Christophe Touret, Corinne Wong, Carleigh Wood, Bokai Yu, Pierre Andrews, Can Balioglu, Peng-Jen Chen, Marta R. Costa-jussà, Maha Elbayad, Hongyu Gong, Francisco Guzmán, Kevin Heffernan, Somya Jain, Justine Kao, Ann Lee, Xutai Ma, Alex Mourachko, Benjamin Peloquin, Juan Pino, Sravya Popuri, Christophe Ropers, Safiyyah Saleem, Holger Schwenk, Anna Sun, Paden Tomasello, Changhan Wang, Jeff Wang, Skyler Wang, and Mary Williamson. 2023. Seamless: Multilingual Expressive and Streaming Speech Translation. https://doi.org/10.48550/ARXIV.2312.05187 Version Number: 1.

[6] Erin R Conway and Helen J Chenery. 2016. Evaluating the MESSAGE Communication Strategies in Dementia training for use with community-based aged care staff working with people with dementia: a controlled pretest–post-test study. *Journal of Clinical Nursing* 25, 7-8 (April 2016), 1145–1155. https://doi.org/10.1111/jocn.13134

[7] Dagoberto Cruz-Sandoval, Jesus Favela, and Eduardo B. Sandoval. 2018. Strategies to Facilitate the Acceptance of a Social Robot by People with Dementia. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Chicago IL USA, 95–96. https://doi.org/10.1145/3173386.3177081

[8] Dagoberto Cruz-Sandoval, Arturo Morales-Tellez, Eduardo Benitez Sandoval, and Jesus Favela. 2020. A Social Robot as Therapy Facilitator in Interventions to Deal with Dementia-related Behavioral Symptoms. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Cambridge United Kingdom, 161–169. https://doi.org/10.1145/3319502.3374840

[9] Emma Dixon and Amanda Lazar. 2020. Approach Matters: Linking Practitioner Approaches to Technology Design for People with Dementia. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–15. https://doi.org/10.1145/3313831.3376432

[10] Julen Etxaniz, Gorka Azkune, Aitor Soroa, Oier Lopez de Lacalle, and Mikel Artetxe. 2023. Do Multilingual Language Models Think Better in English? https://doi.org/10.48550/ARXIV.2308.01223 Version Number: 1.

[11] Jesse Hoey, Krists Zutis, Valerie Leuty, and Alex Mihailidis. 2010. A tool to promote prolonged engagement in art therapy: design and development from arts therapist requirements. In *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*. ACM, Orlando Florida USA, 211–218. https://doi.org/10.1145/1878803.1878841

[12] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685* (2021).

[13] D. Jurafsky, R. Bates, N. Coccaro, R. Martin, M. Meteer, K. Ries, E. Shriberg, A. Stolcke, P. Taylor, and C. Van Ess-Dykema. 1997. Automatic detection of discourse structure for speech recognition and understanding. In *1997 IEEE Workshop on Automatic Speech Recognition and Understanding Proceedings*. IEEE, Santa Barbara, CA, USA, 88–95. https://doi.org/10.1109/ASRU.1997.658992

[14] Keita Kiuchi, Kouyou Otsu, and Yugo Hayashi. 2023. Psychological insights into the research and practice of embodied conversational agents, chatbots and social assistive robots: a systematic meta-review. *Behaviour & Information Technology* (Nov. 2023), 1–41. https://doi.org/10.1080/0144929X.2023.2286528

[15] Peter Klein and Martina Uhlig. 2016. Interactive Memories: technology-aided reminiscence therapy for people with dementia. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, Corfu Island Greece, 1–2. https://doi.org/10.1145/2910674.2935838

[16] Noriaki Kuwahara, Shinji Abe, Kiyoshi Yasuda, and Kazuhiro Kuwabara. 2006. Networked reminiscence therapy for individuals with dementia by using photo and video sharing. In *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*. ACM, Portland Oregon USA, 125–132. https://doi.org/10.1145/1168987.1169010

[17] Fong Yoke Leng, Donald Yeo, Stacey George, and Christopher Barr. 2014. Comparison of iPad applications with traditional activities using person-centred care approach: Impact on well-being for persons with dementia. *Dementia* 13, 2 (March 2014), 265–273. https://doi.org/10.1177/1471301213494514

[18] Marie-Conception Leocadie, Jean-Manuel Morvillers, Sophie Pautex, and Monique Rothan-Tondeur. 2020. Characteristics of the skills of caregivers of people with dementia: observational study. *BMC Family Practice* 21, 1 (Dec. 2020), 149. https://doi.org/10.1186/s12875-020-01218-6

[19] Bhavitvya Malik. 2024. bhavitvyamalik/DialogTag. https://github.com/bhavitvyamalik/DialogTag

[20] M. Mast, R. Kompe, S. Harbeck, A. Kiessling, H. Niemann, E. Noth, E.G. Schukat-Talamazzini, and V. Warnke. 1996. Dialog act classification with the help of prosody. In *Proceeding of Fourth International Conference on Spoken Language Processing. ICSLP '96*, Vol. 3. IEEE, Philadelphia, PA, USA, 1732–1735. https://doi.org/10.1109/ICSLP.1996.607962

[21] Deanne J. O'Rourke, Michelle M. Lobchuk, Genevieve N. Thompson, and Christina Lengyel. 2022. Connecting Through Conversation: A Novel Video-Feedback Intervention to Enhance Long-Term Care Aides' Person-Centred Dementia Communication. *Gerontology and Geriatric Medicine* 8 (April 2022), 233372142211012. https://doi.org/10.1177/23337214221101266

[22] Kyongok Park, Seonhye Lee, JeongEun Yang, Taekwon Song, and Gwi-Ryung Son Hong. 2019. A systematic review and meta-analysis on the effect of reminiscence therapy for people with dementia. *International Psychogeriatrics* 31, 11 (Nov. 2019), 1581–1597. https://doi.org/10.1017/S1041610218002168

[23] Alexis Plaquet and Hervé Bredin. 2023. Powerset multi-class cross entropy loss for neural speaker diarization. In *INTERSPEECH 2023*. ISCA, 3222–3226. https://doi.org/10.21437/Interspeech.2023-205

[24] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2023. Robust speech recognition via large-scale weak supervision. In *Proceedings of the 40th International Conference on Machine Learning (ICML'23)*. JMLR.org. Place: Honolulu, Hawaii, USA.

[25] Margherita Rampioni, Vera Stara, Elisa Felici, Lorena Rossi, and Susy Paolini. 2021. Embodied Conversational Agents for Patients With Dementia: Thematic Literature Analysis. *JMIR mHealth and uHealth* 9, 7 (July 2021), e25381. https://doi.org/10.2196/25381

[26] Marc Rousseaux, Amandine Sève, Marion Vallet, Florence Pasquier, and Marie Anne Mackowiak-Cordoliani. 2010. An analysis of communication in conversation in patients with dementia. *Neuropsychologia* 48, 13 (Nov. 2010), 3884–3890. https://doi.org/10.1016/j.neuropsychologia.2010.09.026

[27] Urška Smrke, Nejc Plohl, and Izidor Mlakar. 2022. Aging Adults' Motivation to Use Embodied Conversational Agents in Instrumental Activities of Daily Living:

Results of Latent Profile Analysis. *International Journal of Environmental Research and Public Health* 19, 4 (Feb. 2022), 2373. https://doi.org/10.3390/ijerph19042373

[28] Vera Stara, Benjamin Vera, Daniel Bolliger, Susy Paolini, Michiel De Jong, Elisa Felici, Stephanie Koenderink, Lorena Rossi, Viviane Von Doellen, and Mirko Di Rosa. 2021. Toward the Integration of Technology-Based Interventions in the Care Pathway for People with Dementia: A Cross-National Study. *International Journal of Environmental Research and Public Health* 18, 19 (Oct. 2021), 10405. https://doi.org/10.3390/ijerph181910405

[29] Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational Linguistics* 26, 3 (2000), 339–374. https://aclanthology.org/J00-3003 Place: Cambridge, MA Publisher: MIT Press.

[30] Bob Woods, Aimee E Spector, Catherine A Jones, Martin Orrell, and Stephen P Davies. 2005. Reminiscence therapy for dementia. *Cochrane Database of Systematic Reviews* (April 2005). https://doi.org/10.1002/14651858.CD001120.pub2

[31] Danni Yu, Luyang Li, Hang Su, and Matteo Fuoli. 2024. Assessing the potential of LLM-assisted annotation for corpus-based pragmatics and discourse analysis: The case of apology. *International Journal of Corpus Linguistics* (June 2024). https://doi.org/10.1075/ijcl.23087.yu

[32] Tamara Zubatiy, Niharika Mathur, Larry Heck, Kayci L. Vickers, Agata Rozga, and Elizabeth D. Mynatt. 2023. "I don't know how to help with that" - Learning from Limitations of Modern Conversational Agent Systems in Caregiving Networks. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (Sept. 2023), 1–28. https://doi.org/10.1145/3610170