Learning Language Representations for Voice-based Conversational Agents for Older Adults

Khalil Mrini, Chen Chen, Co-PI: Ndapa Nakashole, Co-PI: Nadir Weibel
Computer Science and Engineering, University of California San Diego, La Jolla, CA
{khalil,chenchen,nnakashole,weibel}@ucsd.edu

1 INTRODUCTION

Aging is one of the most significant social transformations in the 21st century [24] and enhancing the Quality of Life (QoL) of the aging population is crucial. In the US, around 49.2 million people are 65 or older, and by 2035, this number will increase to 78 million, projecting older adults to outnumber children for the first time [3]. Thus, there is a growing interest in technology catering to the well-being of older adults [27, 35] and their interactions with voice-based conversational agents [6, 29, 34]. For example, the EU has approved over €5M in 2018 for the NESTORE project: an e-coach which is a mobile app and a wearable device to monitor the health of older adults [9]. This project is undergoing testing and plans to use a chatbot to understand simple user instructions [10].

In this innovation proposal, we propose to **build a voice-based conversational agent for older adults that promotes healthy aging, and provides understandable and explainable answers to health-related questions**. Our proposed work is a continuous effort to our existing research [14].

2 PROPOSED WORK

Our proposed approach aims to tackle weaknesses of current Natural Language Processing (NLP) tools to serve older adults, a marginalized but growing community. To accomplish this aim, we outline four main research directions.

First, instead of requiring older adults to adapt to the speech understood by voice assistants, we propose methods that adapt to and better understand everyday speech. *Second*, we propose to make our voice assistant generate understandable replies, by adopting the personal style of speech of its users. *Third*, we aim to develop models that can provide faithful rationales and explanations to their replies, so that older adults can understand the reasoning of the voice assistant. *Finally*, it is very important for us that our voice assistant satisfy its target audience: our multidisciplinary team plans to run Human-Computer Interaction (HCI) studies with older adults to evaluate our voice assistant in real-world settings.

2.1 Domain Adaptation for Everyday Speech

Studies find that one major challenge towards higher user retention of voice-based conversational agents for older adults is the rigidity of query commands [4, 29]. *How can we enable conversational agents to understand everyday speech to increase user retention among older adults?*

Our prior work [23] shows that the popular transformer architecture can leverage tree structures in formally written text through syntactic probing tasks, yet fail to do so in text written like everyday speech. Existing tree-structured transformers [25, 32] show state-of-the-art results on a wide range of tasks including sentiment analysis and text classification. However, their datasets are not user-written, but rather written in a formal, encyclopedia-like style. We introduce a novel tree-structured transformer architecture. Our method achieves state-of-the-art results in two widely used question answering (QA) benchmark datasets, but not in community question answering datasets, where text is user-written, long and informal. Through probing tasks, we show that not absorbing tree structure information leads to no increase in performance in QA. We hereby demonstrate a weakness in a popular NLU architecture to generalize to everyday speech.

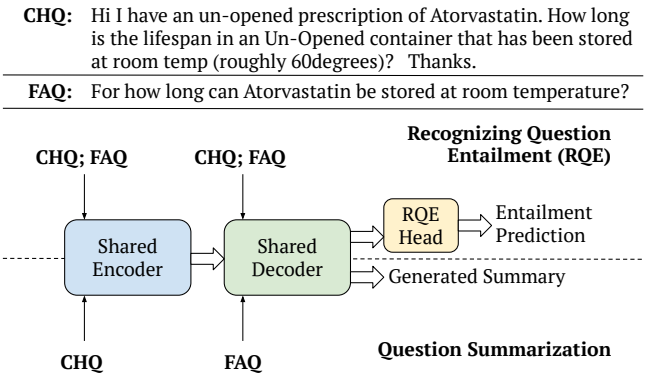

Figure 1: Rather than constraining the speech of older adults, we propose to translate their queries into FAQ-style utterances. Our multi-task learning approach proposed in [20] combines the two tasks of summarization and RQE. We show an example medical question pair, where the first utterance is a Consumer Health Question (CHQ).

For Natural Language Understanding (NLU) to generalize to user queries, we propose, in another work of our own [20], a question understanding approach that "translates" a user query into a formal question. Our approach is to augment datasets to cover both question summarization and recognizing question entailment, and then to train fully shared parameters using a simple multi-task loss objective combining both tasks. We show an overview of our multi-task learning model and an example of a consumer health question and an FAQ-style summary in Figure 1. We compare our results to single-task training on BART, the state-of-the-art in abstractive summarization [17]. We show across 4 medical datasets that our approach is efficient in low-resource settings and performs better than the BART baseline in at least one of human evaluation or ROUGE scores.

In a follow-up project [22], we propose a weighted multi-task loss objective, and a novel gradually soft parameter-sharing approach. In our novel approach, we train one fully shared encoder and two

gradually soft-shared decoders. We add a loss term to constrain the 12 layers of the decoder to be close in representation space, such that the constraint is gradually loosened from the first layer until the last one, which is entirely task-specific. Our approach is a hybrid, encoder-decoder version of soft parameter-sharing [8] and the parameter-sharing configuration of MT-DNN [19], which has achieved state-of-the-art results on the popular GLUE benchmark [31]. We find through experiments that our approach outperforms other parameter-sharing configurations, as well as existing multi-task learning approaches using summarization.

Our next step is to ground question understanding models to a trusted medical knowledge base of answers. The US National Institutes of Health (NIH) released an FAQ-style dataset of medical question-answer pairs, called MedQuAD [1]. We also want to explore how pre-training on a knowledge base can help question understanding and answering systems. A large body of work has explored knowledge bases for question answering [37–39], and recently transformer models for knowledge bases have emerged [2].

We propose to train an end-to-end system that matches a user question to an FAQ and then to an expert-written answer. We first train the system to detect entailment between a user question and one of the FAQs in our dataset of medical question-answer pairs. Each FAQ is attached to multiple answers, and each answer has multiple sentences. In a second step, the model learns to select salient answer sentences given the user question. We propose to use pre-trained models from our prior work [23] to create semi-supervised labels, and thus create a large medical answer sentence selection dataset. The user will then be provided with an expert-written answer, and can also compare their own question to the matched FAQ.

2.2 Personalized Style Transfer

Previous work [1] mentions "*lexical heterogeneity*" as one of the reasons for which existing NLP systems have problems understanding questions asked by patients. We argue that this difference in vocabulary is likewise an obstacle for patients to understand the replies of NLP systems.

Whereas our initial work focuses on understanding patient questions, we want in a second phase to enable the voice assistant to adopt the speech style of older adults when generating answers. *How can we build a conversational agent that not only speaks to older adults, but also speaks like them?*

We plan to explore novel architectures and methods for conversational agents to adapt a style similar to that of their users when generating responses, while preserving the content of their answer. Applications of style transfer vary and can include for instance Shakespearean-to-modern English style transfer using parallel text [13]. Unpaired and unsupervised text style transfer is more difficult as it requires disentangling style and content [7].

In our project, we are particularly interested in non-parallel text style transfer. Such methods include unsupervised text style transfer using adversarial training [26], and using language models as discriminators to disentangle style from content [36]. The latter method achieves state-of-the-art results on the word substitution decipherment task, and the sentiment manipulation task – which is turning a sentence written in a positive sentiment onto a sentence written in a negative sentiment, and vice-versa.

As part of our initial work [20, 22], we have collected a large dataset of long user-written utterances from online health forums. We also have access to formally worded text data related to health, released by the US NIH. Our first experiments would be to devise methods to transfer the style of formal health-related text to online health forums. The goal is that the voice assistant can provide responses in the same personal vocabulary of the user. We propose to develop a non-parallel text style transfer method that learns continuously as the user speaks, and that is able to adapt using only very few training instances.

We hypothesize that a voice assistant that speaks like its target audience can increase user engagement and retention. As part of the evaluation of our method, we will launch real-world experiments with older adults to measure whether personalized style transfer generates more understandable replies. We will also ask whether these replies make the voice assistant sound more familiar, fluent and human-like.

2.3 Explainable Text Generation and Faithful Rationales

There is growing interest in the interpretability of NLP models. An interpretable framework for relation prediction between clinical entities, *e.g.,* diseases and medication, shows that medical experts benefit from the model's rationales during clinical decision-making [33]. Our project aims to address the lack of such rationales to support older adults. Whereas there are many health literacy screening tools [18], there are few chatbots to support health literacy or people with low health literacy [15]. *How can we train question-answering models for older adults to provide explainable replies?*

Transformers have now become ubiquitous in NLP models, but remain hard to interpret due to their numerous attention distributions. One way to show what transformers have learned is through probing tasks. In our prior work on tree-structured transformers [23], we demonstrate that syntactic structure is not well absorbed for online health forum datasets through probing tasks. These are simple classification tasks where we train a one-layer feed-forward network on top of learned, frozen text representations. Our two probing tasks are: predicting the depth of the syntax tree of a given sentence vector, and predicting the syntactic categories of the constituent spans at the second level of the syntax tree.

In our prior work on interpretable neural syntactic parsing [21], we devise a novel form of self-attention – the Label Attention Layer – where attention heads represent labels. Our novel architecture enables to match attention heads to specific syntactic categories. Using this feature, we are able to uncover that our syntactic parser learns simple linguistic rules like: a sentence is made out of a noun phrase and a verb phrase. Our work introduces a novel direction in interpretable representation learning, as it is the first state-of-the-art parser that explicitly includes interpretability in its model architecture.

In the pursuit of our goal to provide explainable replies to older adults, we will continue to incorporate interpretability in our future

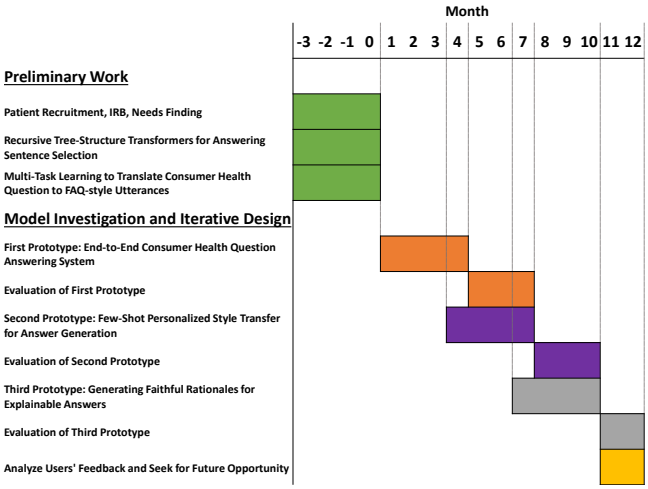

Figure 2: Timeline for proposed research spanning a period of 12 months.

model architectures and to probe what our models learn. One example of a research direction we want to pursue is the generation of textual rationales. A study on Natural Language Inference [16] proposes faithful textual rationales while achieving minimal accuracy loss compared to the state of the art on the SNLI dataset. For us to explore this method, we would potentially create a dataset of human-generated textual rationales for existing medical question-answer pairs. Given that this would be costly, we would like to pursue semi-supervised and weakly supervised ways of creating rationale dataset. One way of doing that would be to use common-sense knowledge transformers (COMET) [2] and explanation templates.

2.4 Real-World Evaluation by Older Adults

A practical inclusive technology design for older adults is important. We propose to integrate and deploy our framework among real-world aging individuals (aged 65 or older), recruited from UC San Diego Outpatient Geriatric Primary Care Clinic.

The design and evaluations of existing voice-based conversational agents for enhancing healthcare management and quality of life only focused on general users, and relied on the training data collected primarily from younger adults [1]. In contrast, understanding voice interaction experience among aging individuals are challenging [30]. This is mainly caused by the degradation of their short-term memory, and results in barriers caused by their unique mental models [11]. In this work, we aim to bridge this gap by integrating our system with existing health care infrastructure, and deploying on the commercially available standalone smart speakers, placed in older adults' residential houses (stand-alone home and retirement community).

As part of our preliminary work, we have conducted large-scale semi-structured interviews with providers and patients (aged 68 to 90), and identified the key barriers that older adults might encounter during daily life and while managing their health [4]. Our study protocol has been approved by the Institutional Review Board at UC San Diego.

To this end, based on our preliminary work and proposed framework, our proposed contributions are two-fold. (1) We will contribute full-fledged evaluations of the effectiveness of proposed framework. *Qualitatively*, we will evaluate the user experience by obtaining subjective comments, which can provide user feedback regarding fluency, coherence and comprehensibility of generated responses. *Quantitatively*, we will also capture the underlying metrics during the interactions between conversational agents and older adults, *e.g.,* users' response time and acknowledge intent regarding particular generated responses [12]. (2) We will also seek opportunity to contribute open-sourced and de-identified interaction dataset that would be potentially useful for researchers and practitioners.

3 RESEARCH PLAN

Figure 2 shows an horizontal view of our proposed study timeline. We target top-tier conferences organized by the ACL (like ACL, EMNLP) and the SIGCHI (like CHI).

4 TEAM, ADVISORS AND COLLABORATORS

This proposed idea is a continuous effort of our ongoing project: *VOLI*, or voice assistant for quality of life and healthcare improvement in aging populations. For more details about our project, please visit our website: http://voli.ucsd.edu. The faculty advisors of our proposed project include Ndapa Nakashole, Ph.D. and Nadir Weibel, Ph.D. from Department of Computer Science and Engineering at UC San Diego, Emilia Farcas, Ph.D. from Qualcomm Institute, the UC San Diego Division of California Institute for Telecommunications and Information Technology (Calit2). Our faculty advisors also include clinical experts affiliated with UC San Diego Health, one of the top academic health system in the Nation, including Alison Moore, M.D., M.PH. as well as Michael Hogarth, M.D.

Khalil Mrini is a third year PhD Student in Computer Science at UC San Diego, advised by Prof. Ndapa Nakashole. His PhD thesis is on Interpretability and Generalization in Applications of Text Representation Learning. He is developing interpretable neural models, studying what they learn, and making them generalizable to a wider range of applications and environments. He has previously published work on an interpretable neural syntactic parser, which set a new state of the art in English and Chinese benchmark datasets [21]. His most recent projects show how tree-structured representations may not understand user queries [23], and propose to alleviate this problem by translating user queries into FAQs with matched answers [20, 22]. He has previously interned at Adobe Research and Amazon Alexa. For more information about him, please visit https://khalilmrini.github.io.

Chen Chen is a third year PhD Student in Computer Science and Engineering at UC San Diego, advised by Prof. Nadir Weibel. His research interest is on accessibility, human computer interaction and mobile computing. He is in particular interested in exploring novel interaction paradigms for minorities, and solving situational as well as temporal disabilities for majorities. As an outstanding HCI innovator, his previous works were published in several top tier conferences including ACM SenSys 2020 [28] and IPSN 2021 [5]. He has been recognized with Best Poster Award in ACM SenSys 2020 and Qualcomm Innovation Fellowship Finalist in 2019. For more information about him, please visit https://chen-chen.me.

We are confident that our multidisciplinary team of NLP researchers, HCI researchers and medical doctors has the skill set to push our proposal forward to fruition.

REFERENCES

[1] Asma Ben Abacha and Dina Demner-Fushman. 2019. A question-entailment approach to question answering. *BMC bioinformatics* 20, 1 (2019), 511.

[2] Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. 2019. COMET: Commonsense Transformers for Automatic Knowledge Graph Construction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 4762–4779.

[3] United States Census Bureau. 2018. Older People Projected to Outnumber Children for First Time in U.S. History. (03 2018). https://www.census.gov/newsroom/press-releases/2018/cb18-41-population-projections.html

[4] Chen Chen, Janet G Johnson, Kemeberley Charles, ALice Lee, Michael Hogarth, Alison A Moore, Emilia Farcas, and Nadir Weibel. 2021. Alexa is Not Just a Toy: Exploring Needs and Design for Intelligent Voice Assistants and EHR Integration for Older Adults. *Submitted to the 24th ACM Conference on Computer-Supported Cooperative Work and Social Computing* (2021).

[5] Chen Chen, Ke Sun, and Xinyu Zhang. 2021. ExGSense: Toward Facial Gesture Sensing and Reconstruction with a Sparse Near-Eye Sensor Array. In *Proceedings of the 2021 International Conference on Information Processing in Sensor Networks* (Nashville, Tennessee, USA) *(IPSN '19)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/ToAppear

[6] Kyungjin Chung, Young Hoon Oh, and Da Young Ju. 2019. Elderly Users' Interaction with Conversational Agent. In *Proceedings of the 7th International Conference on Human-Agent Interaction*. 277–279.

[7] Ning Dai, Jianze Liang, Xipeng Qiu, and Xuan-Jing Huang. 2019. Style Transformer: Unpaired Text Style Transfer without Disentangled Latent Representation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 5997–6007.

[8] Long Duong, Trevor Cohn, Steven Bird, and Paul Cook. 2015. Low resource dependency parsing: Cross-lingual parameter sharing in a neural network parser. In *Proceedings of the 53rd annual meeting of the Association for Computational Linguistics and the 7th international joint conference on natural language processing (volume 2: short papers)*. 845–850.

[9] Mira El Kamali, Leonardo Angelini, Maurizio Caon, Giuseppe Andreoni, Omar Abou Khaled, and Elena Mugellini. 2018. Towards the nestore e-coach: a tangible and embodied conversational agent for older adults. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 1656–1663.

[10] Mira El Kamali, Leonardo Angelini, Maurizio Caon, Denis Lalanne, Omar Abou Khaled, and Elena Mugellini. 2020. An embodied and ubiquitous E-coach for accompanying older adults towards a Better lifestyle. In *International Conference on Human-Computer Interaction*. Springer, 23–35.

[11] Arthur Fisk, Arthur D., Wendy Rogers, Wendy A., Neil Charness, Neil, Sara Czaja, John Sara, Sharit, and Joseph. 2004. *Designing for Older Adults: Principles and Creative Human Factors Approaches*. https://doi.org/10.1201/9781420023862

[12] Seyyed Hadi Hashemi, Kyle Williams, Ahmed El Kholy, Imed Zitouni, and Paul A. Crook. 2018. Measuring User Satisfaction on Smart Speaker Intelligent Assistants Using Intent Sensitive Query Embeddings. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management* (Torino, Italy) *(CIKM '18)*. Association for Computing Machinery, New York, NY, USA, 1183–1192. https://doi.org/10.1145/3269206.3271802

[13] Harsh Jhamtani, Varun Gangal, Eduard Hovy, and Eric Nyberg. 2017. Shakespearizing Modern Language Using Copy-Enriched Sequence to Sequence Models. In *Proceedings of the Workshop on Stylistic Variation*. 10–19.

[14] Janet G. Johnson, Khalil Mrini, Michael Hogarth, Alison A Moore, Ndapa Nakashole, Nadir Weibel, and Emilia Farcas. 2020. Voice-Based Conversational Agents for Older Adults. In *Proceedings of the CHI 2020 Workshop on Conversational Agents for Health and Wellbeing*. Association for Computing Machinery, Honolulu, Hawaii.

[15] Rafal Kocielnik, Elena Agapie, Alexander Argyle, Dennis T Hsieh, Kabir Yadav, Breena Taira, and Gary Hsieh. 2019. HarborBot: A Chatbot for Social Needs Screening. In *AMIA Annual Symposium Proceedings*, Vol. 2019. American Medical Informatics Association, 552.

[16] Sawan Kumar and Partha Talukdar. 2020. NILE : Natural Language Inference with Faithful Natural Language Explanations. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Online, 8730–8742. https://doi.org/10.18653/v1/2020.acl-main.771

[17] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 7871–7880.

[18] Hongyan Liu, Huan Zeng, Yang Shen, Fan Zhang, Manoj Sharma, Weiyun Lai, Yu Zhao, Genhui Tao, Jun Yuan, and Yong Zhao. 2018. Assessment tools for health literacy among the general population: a systematic review. *International journal of environmental research and public health* 15, 8 (2018), 1711.

[19] Xiaodong Liu, Pengcheng He, Weizhu Chen, and Jianfeng Gao. 2019. Multi-Task Deep Neural Networks for Natural Language Understanding. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 4487–4496.

[20] Khalil Mrini, Franck Dernoncourt, Walter Chang, Emilia Farcas, and Ndapa Nakashole. 2021. Towards Improving Abstractive Summarization and Entailment Recognition for Consumer Health Questions. *Submitted, under review* (2021).

[21] Khalil Mrini, Franck Dernoncourt, Quan Hung Tran, Trung Bui, Walter Chang, and Ndapandula Nakashole. 2020. Rethinking self-attention: Towards interpretability in neural parsing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings*. 731–742.

[22] Khalil Mrini, Franck Dernoncourt, Seunghyun Yoon, Trung Bui, Walter Chang, Emilia Farcas, and Ndapa Nakashole. 2021. A Gradually Soft Multi-Task and Data-Augmented Approach
to Medical Question Understanding. *Submitted, under review* (2021).

[23] Khalil Mrini, Emilia Farcas, and Ndapa Nakashole. 2021. Recursive Tree-Structured Self-Attention for Answer Sentence Selection. *Submitted, Under Review* (2021).

[24] United Nations. 2020. World Population Ageing 2019 – Highlights. (2020). https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf

[25] Xuan-Phi Nguyen, Shafiq Joty, Steven Hoi, and Richard Socher. 2019. Tree-Structured Attention with Hierarchical Accumulation. In *International Conference on Learning Representations*.

[26] Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi S Jaakkola. 2017. Style Transfer from Non-Parallel Text by Cross-Alignment. In *Neural Information Processing Systems*.

[27] James Simpson, Franziska Gaiser, Miroslav Macík, and Timna Breßgott. 2020. Daisy: A Friendly Conversational Agent for Older Adults. In *Proceedings of the 2nd Conference on Conversational User Interfaces*. 1–3.

[28] Ke Sun, Chen Chen, and Xinyu Zhang. 2020. "Alexa, Stop Spying on Me!": Speech Privacy Protection against Voice Assistants. In *Proceedings of the 18th Conference on Embedded Networked Sensor Systems*. Association for Computing Machinery, New York, NY, USA, 298–311. https://doi.org/10.1145/3384419.3430727

[29] Milka Trajkova and Aqueasha Martin-Hammond. 2020. "Alexa is a Toy": Exploring Older Adults' Reasons for Using, Limiting, and Abandoning Echo *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376760

[30] John Vines, Gary Pritchard, Peter Wright, Patrick Olivier, and Katie Brittain. 2015. An Age-Old Problem: Examining the Discourses of Ageing in HCI and Strategies for Future Research. *ACM Trans. Comput.-Hum. Interact.* 22, 1, Article 2 (Feb. 2015), 27 pages. https://doi.org/10.1145/2696867

[31] Alex Wang, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel Bowman. 2018. GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding. In *Proceedings of the 2018 EMNLP Workshop BlackboxNLP: Analyzing and Interpreting Neural Networks for NLP*. 353–355.

[32] Yaushian Wang, Hung-Yi Lee, and Yun-Nung Chen. 2019. Tree Transformer: Integrating Tree Structures into Self-Attention. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. 1060–1070.

[33] Zhen Wang, Jennifer Lee, Simon Lin, and Huan Sun. 2020. Rationalizing Medical Relation Prediction from Corpus-level Statistics. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.

[34] Pierre Wargnier, Giovanni Carletti, Yann Laurent-Corniquet, Samuel Benveniste, Pierre Jouvelot, and Anne-Sophie Rigaud. 2016. Field evaluation with cognitively-impaired older adults of attention management in the embodied conversational agent louise. In *2016 IEEE International Conference on Serious Games and Applications for Health (SeGAH)*. IEEE, 1–8.

[35] Nirmalie Wiratunga, Kay Cooper, Anjana Wijekoon, Chamath Palihawadana, Vanessa Mendham, Ehud Reiter, and Kyle Martin. 2020. FitChat: Conversational Artificial Intelligence Interventions for Encouraging Physical Activity in Older Adults. *arXiv preprint arXiv:2004.14067* (2020).

[36] Zichao Yang, Zhiting Hu, Chris Dyer, Eric P Xing, and Taylor Berg-Kirkpatrick. 2018. Unsupervised text style transfer using language models as discriminators. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*. 7298–7309.

[37] Scott Wen-tau Yih, Ming-Wei Chang, Xiaodong He, and Jianfeng Gao. 2015. Semantic parsing via staged query graph generation: Question answering with knowledge base. (2015).

[38] Wen-tau Yih, Matthew Richardson, Christopher Meek, Ming-Wei Chang, and Jina Suh. 2016. The value of semantic parse labeling for knowledge base question answering. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 201–206.

[39] Mo Yu, Wenpeng Yin, Kazi Saidul Hasan, Cicero dos Santos, Bing Xiang, and Bowen Zhou. 2017. Improved neural relation detection for knowledge base question answering. *arXiv preprint arXiv:1704.06194* (2017).