Eliciting User Self-disclosure using Reciprocity in Human-Voicebot Conversations

Wieke Noa Harmsen*
Iris Hendrickx
wieke.harmsen@ru.nl
iris.hendrickx@ru.nl
Centre for Language Studies,Centre for Language and Speech Technology, Radboud University
Nijmegen, The Netherlands

Jelte van Waterschoot*
Mariët Theune
j.b.vanwaterschoot@utwente.nl
m.theune@utwente.nl
Human Media Interaction, University of Twente
Enschede, The Netherlands

ABSTRACT

In this study, we have developed a voicebot which asks users questions about their daily activities and social participation to gain insights into their happiness and well-being. We hypothesize that showing disclosure when asking questions can elicit reciprocity of self-disclosure by the users. We define two types of disclosure: self-disclosure and other-disclosure. Self-disclosure is sharing thoughts, feelings and information about oneself, whereas other-disclosure is sharing information about others and opinions of others. We analyzed 122 answers to the voicebot's disclosure and control questions by annotating the number of self-disclosure statements in the answers. We found no significant effect of asking disclosure questions on the number of self-disclosure statements. However, we did find a positive effect of asking disclosure questions on common markers of reciprocity such as the number of words, topic phrases, and first-person pronouns. Replication of this study with more participants would strengthen the validity of the findings.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in interaction design**; *Sound-based input / output*; *Web-based interaction.*

KEYWORDS

Conversational Agents, Self-Disclosure, Automatic Speech Recognition, Well-being

ACM Reference Format:
Wieke Noa Harmsen, Iris Hendrickx, Jelte van Waterschoot, and Mariët Theune. 2023. Eliciting User Self-disclosure using Reciprocity in Human-Voicebot Conversations. In *ACM conference on Conversational User Interfaces (CUI '23), July 19–21, 2023, Eindhoven, Netherlands.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3571884.3604301

1 INTRODUCTION

In this study, we investigate how conversations with speech-based chatbots (voicebots) can be used to gain insights into people's happiness and well-being, in order to improve them. This study is part of the Behaviour-based Language-Interactive Speaking Systems (BLISS) project [23]. Since sharing thoughts with someone else relieves stress [11], talking about yourself is good for your mental health. However, revealing personal information can make one feel vulnerable, and for that reason an anonymous chat with a voicebot can be easier than a direct conversation with a person [13].

We approach the measurement of happiness from a positive health perspective [9]. This means that we would like to learn about someone's general well-being, and not about specific happiness moments. Therefore, the voicebot asks questions about the user's daily activities and social participation. To be able to extract information about happiness and well-being from the users' answers, it is important that the users give answers in which they reveal personal information. In other words: they need to disclose information about their thoughts and feelings. This is also called *self-disclosure*. Self-disclosure has been found to positively influence mental and physical health [4, 15] and enhance the positive effect of emotional support on worry reduction [14]. In addition, disclosing to a person or chatbot has comparable psychological impact [7].

Earlier research has shown that self-disclosure of one conversation partner induces self-disclosure of the other, both in human-human [2, 3] and human-machine interactions [12, 16, 17]. This effect is called reciprocity of self-disclosure.

In the current study, we investigate whether the self-disclosure reciprocity effect is present in the type of human-machine interaction where a voicebot asks the user questions about their social participation and daily life to gain insight into their happiness and well-being. We chose speech as the main means of interaction to increase the accessibility for older people, who have difficulty reading or typing [1]. In addition, speech is a faster and more efficient input method than text [19, 20].

We define disclosure by the user as any voluntarily shared statement that is not required to answer the question [16]. Table 2 shows an example. We distinguish between *self-disclosure* and *other-disclosure*. Self-disclosure involves disclosure of information about oneself such as thoughts, feelings and opinions, whereas other-disclosure is information about others' thoughts, feelings and opinions, which is still relevant to a person's environment. For a voicebot, it may be more fitting to show other-disclosure than

*Both authors contributed equally to this research.

CUI '23, July 19–21, 2023, Eindhoven, Netherlands

ACM ISBN 979-8-4007-0014-9/23/07.
https://doi.org/10.1145/3571884.3604301

self-disclosure, because the voicebot does not have thoughts or feelings of its own.

Our research questions are as follows:

RQ1: *To what extent can a voicebot showing disclosure elicit disclosure by the users?*

RQ2: *Is there a difference in the amount of disclosure elicited by a voicebot showing self-disclosure or other-disclosure?*

We hypothesize that the use of other-disclosure while asking the questions will elicit at least the same amount of disclosure by the users as self-disclosure, since other-disclosure gives examples of self-disclosure by others and is assumed to be more believable than self-disclosure by an information gathering conversational agent.

2 QUESTION DESIGN

The conversations with the voicebot aim to collect information about the user's general happiness and well-being. We approach this from a positive health perspective and use a broad happiness model that has six dimensions: quality of life, bodily and mental functions, social participation, spiritual dimension and daily functioning [9]. Our study focuses on collecting information from the users on two of these dimensions: daily functioning and social participation. We chose these dimensions because a future version of the voicebot should be able to suggest new social or daily activities that would fit the interests of the user to increase their social network and healthy behavior. In addition, it is possible to ask very concrete questions about these two dimensions of happiness, which anyone can talk about. This is important, since we learned from earlier prototypes of our voicebot that users found questions about happiness often abstract and therefore difficult to answer.

We used existing questionnaires about daily life activities and social participation [18] to formulate the voicebot's questions. This resulted in a conversation of five blocks of questions. In each block, all questions are about the same topic. The topics are: family, computer games, solitaire games, board games and pets. The first question of a question block is the disclosure question or control question. In both conditions, the questions are the same. However, in the disclosure condition, the question is introduced with extra information in which the voicebot shows disclosure.

We used two strategies to create the disclosure questions. In the first strategy, the voicebot says something about what it thinks or did. We refer to this type of disclosure as *fictional self-disclosure.* This information is always a lie, since the voicebot does not have thoughts or feelings of its own. For example:

"As a computer, I often play games against myself. My great example is my predecessor DeepBlue. You know, the chess program that beats chess grandmaster Kasparov. Do you often play games on your own, like patience, making jigsaw puzzles, or doing crossword puzzles?"

In the second strategy, the voicebot shares information about information elicited in (fictional) previous conversations with others. These questions start for example with *"Recently, I spoke to someone who thought ..."* or *"Others often mention that..."*. These statements could in principle be true. However, in this study, they are all scripted. We refer to this type of disclosure as *other-disclosure.*

After asking the initial question in a topic block, some follow-up questions are asked, based on the answers of the user. These questions do not contain self-disclosure or disclosure of others. For a complete overview of all topics and the corresponding disclosure and control questions, we refer to Table 1.

3 METHOD

Below we describe our implementation, participants, the set-up of our study and the evaluation. The voicebot communicates in Dutch and the participants were all fluent Dutch speakers.

3.1 Implementation

We modified a web browser-based chatbot from Games for Health[1] for speech-based interaction. We used the KaldiNL speech recognizer [22] on a server for online automatic speech recognition. In addition, we used a male voice from ReadSpeaker[2] for the text-to-speech of the voicebot. Participants could interact with the voicebot on their tablets, phones or computers. When interacting with the voicebot, users had to press a microphone button once to start recording their speech, and again to stop recording.

The study has a between-subject set-up, so we created two different versions of our question-asking voicebot: the disclosure-voicebot and the control-voicebot. The participants were randomly assigned to one of the two conditions. The disclosure-voicebot always disclosed before asking a question and the control-voicebot directly asked the question (see Table 1).

3.2 Participants

We used our personal networks and social media for recruitment. In total, we collected 64 conversations in six weeks. However, due to technical problems, we could only use data of 34 conversations in the current study.

3.3 Study Setup

Once people clicked on the experiment link, they were first forwarded to a questionnaire, in which they had to give informed consent for participating in the study and provide some personal information like age-range and region of growing up. Next, they were redirected to the experiment. After an audio test to check the microphone and sound, the conversation with the voicebot was started. Each conversation started with a welcome statement, followed by the five question blocks in a random order. The study was ethically approved by the Ethics Assessment Committee Humanities from Radboud University, The Netherlands (case number ETC-GW 2020-9960).

3.4 Analysis of the data

3.4.1 Data selection. We collected the audio recordings and ASR-transcriptions of the users' answers to the questions by the voicebot. We also manually transcribed the audio files and computed the Word Error Rate (WER) of the ASR-transcriptions. We found a high WER of 51.7%. Therefore, we used the manual transcriptions for the rest of the analysis.

We only analysed the relevant answers for this research, which are the direct answers to the disclosure questions and the direct

[1] https://gamesfor.health/

[2] https://www.readspeaker.com/

Table 1: An overview of all disclosure questions and their corresponding control question (translated from Dutch).

Topic	Type	Question
Family	Disclosure (other)	If you think about your family, who is the first person that pops up in your mind? I would like to know a little more about your relationship. Others often mention that someone is patient, exudes calmness, or is always there for them, but what does that person mean to you?
	Control	If you think about your family, who is the first person that pops up in your mind? What does that person mean to you?
Computer games	Disclosure (other)	When I talk to someone who's under thirty, they often answer me that they like to play games on the computer, most of the time with friends. Do you like to do that as well?
	Control	Do you like to play games on the computer?
Solitaire games	Disclosure (self)	As a computer, I often play games against myself. My great example is my predecessor DeepBlue. You know, the chess program that beat chess grandmaster Kasparov. Do you often play games on your own, like patience, making jigsaw puzzles, or doing crossword puzzles?
	Control	Do you often play games on your own, like patience, making jigsaw puzzles, or doing crossword puzzles?
Board Games	Disclosure (other)	Recently, I spoke to someone who told me that she loves playing Wordfeud. That's a kind of digital variant of Scrabble. Do you like playing board games? If yes, which board game do you like to play?
	Control	Do you like playing board games? If yes, which board game do you like to play?
Pets	Disclosure (self)	As Babbelbot, I don't just talk to people, but I also read regularly online articles to stay informed. For example, I recently read in an article that thanks to the Coronacrisis chickens became very popular pets! Do you have pets? If yes, which pets?
	Control	Do you have pets? If yes, which pets?

answers to their corresponding control questions. A direct answer is an answer directly following a disclosure question or its corresponding control question, this is the first question in the topic block. This resulted in two groups:

(1) direct answers to disclosure questions (dir-disc)
(2) direct answers to control questions (dir-cont)

3.4.2 Disclosure annotation. Two authors manually annotated the answers with the number of disclosure statements they contained, following the self-disclosure definition of [16], in which any type of information that is not requested and voluntarily shared is labeled as disclosure. Table 2 shows a fictional example. When a statement was labeled as disclosure, the annotators also specified whether it was self-disclosure or other-disclosure. Self-disclosure involves disclosure of information about oneself, whereas other-disclosure is disclosure of information about others, such as the opinion of someone else. The inter-annotator agreement was substantial ($\kappa(self - disclosure) = 0.68$; $\kappa(other - disclosure) = 0.74$). The annotators discussed their disagreements and came up with one gold-standard.

3.4.3 Reciprocity measures. For each of the direct answers, we then computed four automatic measures which are markers of reciprocity: the answer length in the number of words and in number of topic phrases, the number of singular and plural first-person pronouns (both possessive and personal) and the percentage of word overlap [5, 10, 16]. The number of topic phrases is the total number of verb and noun phrases in an answer. We calculated this using Python's package spaCy [8] using dependency trees and part-of-speech tagging. Since we are using speech data, which contains stalling words (e.g., "uhm"), we expect that the number of topic phrases might be a more robust measure to compute answer length than number of words. We computed the word overlap as the lemma overlap between the question and the answer. For each answer, we normalized the word overlap score by dividing the number of overlapping words by the total number of words in the question [16]. As long as words had the same lemma, they were counted as word overlap.

Agent

Do you often play games on your own, like solitaire or making jigsaw puzzles or crossword puzzles?

User 1 (self-disclosure)

That's something I don't do often on my own, **but playing games with other people is something I do like.**

User 2 (other-disclosure)

No, **but my husband likes making sudoku puzzles in the newspaper.**

Table 2: Two ways of answering the voicebot's question. The bold parts of the answer are non-required and voluntarily shared information and thus disclosure. The answer of user 1 contains a self-disclosure statement. The answer of user 2 contains an other-disclosure statement.

4 RESULTS

In this section we describe our data set, the analysis of the user answers and post-hoc analyses.

4.1 Description of the data

The analysed data consist of 121 answers to disclosure and control questions in 34 different conversations. Table 3 shows an overview of the number of direct answers to disclosure questions (dir-disc) and direct answers to control questions (dir-cont). The dir-disc group can be split in two subgroups: answers to self-disclosure questions (dir-disc-self) and answers to other-disclosure (dir-disc-other) questions.

Group	Subgroup	Nr. of answers
dir-disc		66
	dir-disc-self	*29*
	dir-disc-other	*37*
dir-cont		56
total		122

Table 3: The number of answers given to disclosure questions (dir-disc) and control questions (dir-cont)

4.2 Disclosure measures

We calculated the mean and standard deviation of the number of self-disclosure statements and other-disclosure statements in answers to (self/other) disclosure questions and answers to control questions (Table 4). From the aggregated data, we can conclude that the number of other-disclosure statements in the answers is very low. Other-disclosure questions are never answered with an other-disclosure statement, and the mean number of other-disclosure statements in answers to self-disclosure and control questions is respectively 0.03 and 0.04. Inspection of the answers shows that all other-disclosure statements in our data set are statements in which something is told about a pet, for example its name.

Measure	Group	Mean	Std
# Self-disclosure statements	dir-disc	0.79	0.95
	dir-disc-self	0.83	0.92
	dir-disc-other	0.72	0.99
	dir-cont	0.61	0.86
# Other-disclosure statements	dir-disc	0.02	0.12
	dir-disc-self	0.03	0.18
	dir-disc-other	0	0
	dir-cont	0.04	0.19

Table 4: Aggregated data of the disclosure statements

From Table 4 we can also conclude that the number of self-disclosure statements is higher in answers to disclosure questions (dir-disc, mean = 0.79) than in answer to control questions (dir-cont, mean = 0.61). A Mann-Whitney U-test showed that this difference is not significant ($p > 0.05$). Within the dir-disc group (dir-disc-self vs dir-disc-other), we find that the number of self-disclosure statements is higher in answers to questions in which the bot self-discloses. However, a Mann-Whitney U-test shows that this difference is also not significant ($p > 0.05$).

4.3 Reciprocity measures

For each of the direct answers, the number of words (#words), number of topic phrases (#topics), the number of first-person pronouns (# first-person pronouns) and the word overlap rate (% word overlap) were computed. These are all common markers of reciprocity [5, 10, 16]. The aggregated data of each of these measures for each group are displayed in Table 5. We conducted one-sided Mann-Whitney U-tests to see if there was a statistically significant difference between the scores on these four measures in the disclosure condition and the control condition. We found that the number of words ($p < 0.01$), number of topic phrases ($p < 0.005$), number of first-person pronouns ($p < 0.001$) and word overlap rate ($p < 0.001$) were all four significantly higher in the disclosure condition than in the control condition.

Measure	Group	Mean	Std
# words	dir-disc	11.17	9.89
	dir-cont	9.05	10.27
# topic phrases	dir-disc	2.71	2.27
	dir-cont	1.80	2.12
# first-person pronouns	dir-disc	1.24	1.05
	dir-cont	0.73	0.97
% word overlap	dir-disc	23.39	23.89
	dir-cont	7.29	13.23

Table 5: Aggregated data of the reciprocity measures.

5 CONCLUSION AND DISCUSSION

In this study, we developed a Dutch voicebot that uses self-disclosure and other-disclosure to ask questions. We found that our voicebot did not elicit significantly more disclosure by the user when it asked questions using disclosure than when it asked questions without disclosure. The type of disclosure question (self-disclosure or other-disclosure) did not influence the amount of disclosure elicited. The first finding is not in line with our expectations, since earlier studies have shown that self-disclosure by one conversation partner elicits self-disclosure by the other, both in human-human [2, 3] and human-machine interactions [12, 16, 17].

These findings can probably be explained by the fact that they are based on a small number of conversations with the voicebot and by the limited diversity of self-disclosure and other-disclosure questions. Unfortunately, we only had data from 34 participants in our final data set, answering three other-disclosure and two self-disclosure questions. Another possible explanation is the question design. Currently, the voicebot shows both self-disclosure and other-disclosure while asking the questions, while in previous studies [16, 17] only self-disclosure was used. A third explanation could be the interview style of the conversation. In a more social context, like [16, 17], users might be more likely to share information about themselves with a bot.

Secondly, we found that direct answers to disclosure questions contain significantly more words, more topic phrases, more first-person pronouns and more word overlap with the preceding question than direct answers to the corresponding control questions. This finding is in line with our expectations [16]. Since these are all measures of reciprocity [5, 10, 16], we can conclude from these automatic measures that there is more reciprocity in the disclosure condition than in the control condition. In combination with our findings about disclosure, we have to mention that this does not mean that self-disclosure is reciprocated, as potentially the users were only mirroring/reciprocating the response length of the voicebot [6]. However, the fact that users not only talked more but also mentioned themselves more often in response to disclosure by the bot does suggest that with more data, it might be possible to find such an effect.

6 LIMITATIONS AND FUTURE DIRECTIONS

Here we discuss the limitations of this study and suggest solutions to overcome them. We also mention some future research directions.

A limitation is that the control questions are not exactly a subset of the disclosure questions. Although the meaning is the same, this disparity should be avoided in future studies. A challenge is that we made a distinction between self- and other-disclosure, and found much less in the latter category. A future study should investigate if such a distinction is beneficial to understand disclosure better.

We focused solely on disclosure in this study, but the bots' disclosure questions can also lead to more enjoyable conversations, a factor that was not measured in this experiment. So it would be interesting to add a post-experiment questionnaire asking the participants to what extent they engaged with and trusted the voicebot [21]. Furthermore, asking the users whether they were aware of the fact that everything that the voicebot discloses is fictional could help to formulate better disclosure questions. Additionally, we could ask users to what extent the information they shared with the voicebot was real, and to what extent the answers were made-up. This could reveal insights into how seriously they took the conversation.

Finally, we would like to discuss the measures we used in this study. We annotated disclosure manually in the answers. In a future study, we could see to what extent this could be done automatically, similar to earlier work for the English language [16]. However, this would require a lot more data. In addition, we should take a more refined approach to extracting the topics for measuring reciprocity. For instance, topics that were already mentioned in the question might not be considered as new topics in the answer, because the voicebot primed those topics. A more fair measure of self-disclosure would be to filter these topics and only look at how many new topics are shared without explicitly asking for them.

ACKNOWLEDGMENTS

We would like to thank Games for Health and ReadSpeaker for making their software available for this project. This work is part of the research programme Data2Person with project no. 628.011.029, which is (partially) financed by the Dutch Research Council (NWO).

REFERENCES

[1] Lorainne Tudor Car, Dhakshenya Ardhithy Dhinagaran, Bhone Myint Kyaw, Tobias Kowatsch, Shafiq Joty, Yin-Leng Theng, and Rifat Atun. 2020. Conversational Agents in Health Care: Scoping Review and Conceptual Analysis. *Journal of Medical Internet Research* 22, 8 (2020), e17158. https://doi.org/10.2196/17158

[2] Gordon J. Chelune. 1979. Measuring Openness in Interpersonal Communication. In *Self-Disclosure: Origins, Patterns, and Implications of Openness in Interpersonal Relationships*. Jossey-Bass, San Francisco, CA, USA, 1–27.

[3] Valerian J Derlaga and John H Berg. 1987. *Self-disclosure: Theory, research, and therapy*. Springer Science & Business Media.

[4] Valerian J. Derlega, Barbara A. Winstead, Robin J. Lewis, and James Maddux. 1993. Clients' Responses to Dissatisfaction in Psychotherapy: A Test of Rusbult's Exit-Voice-Loyalty-Neglect Model. *Journal of Social and Clinical Psychology* 12, 3 (1993), 307–318. https://doi.org/10.1521/jscp.1993.12.3.307

[5] Vernon B. Harper and Erika J. Harper. 2006. Understanding Student Self-Disclosure Typology through Blogging. *The Qualitative Report* 11 (2006), 251–261. https://doi.org/10.46743/2160-3715/2006.1673

[6] Jennifer Hill, W. Randolph Ford, and Ingrid G. Farreras. 2015. Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior* 49 (2015), 245–250. https://doi.org/10.1016/j.chb.2015.02.026

[7] Annabell Ho, Jeff Hancock, and Adam S Miner. 2018. Psychological, Relational, and Emotional Effects of Self-Disclosure After Conversations With a Chatbot. *Journal of Communication* 68, 4 (05 2018), 712–733. https://doi.org/10.1093/joc/jqy026

[8] Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. 2020. SpaCy: Industrial-strength Natural Language Processing in Python. (2020). https://doi.org/10.5281/zenodo.1212303

[9] M. Huber, M. van Vliet, M. Giezenberg, B. Winkens, Y. Heerkens, P. C. Dagnelie, and J. A. Knottnerus. 2016. Towards a 'Patient-centred' Operationalisation of the New Dynamic Concept of Health: a Mixed Methods Study. *BMJ Open* 6, 1 (2016), e010091. https://doi.org/10.1136/bmjopen-2015-010091

[10] Sydney M. Jourard and Peggy E. Jaffe. 1970. Influence of an Interviewer's Disclosure on the Self-disclosing Behavior of Interviewees. *Journal of Counseling Psychology* 17 (1970), 252–257. Issue 3. https://doi.org/10.1037/h0029197

[11] Jeffrey H Kahn and Robert M Hessling. 2001. Measuring the tendency to conceal versus disclose psychological distress. *Journal of Social and Clinical Psychology* 20, 1 (2001), 41–65. https://doi.org/10.1521/jscp.20.1.41.22254

[12] Yi-Chieh Lee, Naomi Yamashita, Yun Huang, and Wai Fu. 2020. "I Hear You, I Feel You": Encouraging Deep Self-disclosure through a Chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 1–12. https://doi.org/10.1145/3313831.3376175

[13] Gale M. Lucas, Albert Rizzo, Jonathan Gratch, Stefan Scherer, Giota Stratou, Jill Boberg, and Louis-Philippe Morency. 2017. Reporting Mental Health Symptoms: Breaking Down Barriers to Care with Virtual Human Interviewers. *Frontiers in Robotics and AI* 4 (2017). https://doi.org/10.3389/frobt.2017.00051

[14] Jingbo Meng and Yue (Nancy) Dai. 2021. Emotional Support from AI Chatbots: Should a Supportive Partner Self-Disclose or Not? *Journal of Computer-Mediated Communication* 26, 4 (05 2021), 207–222. https://doi.org/10.1093/jcmc/zmab005

[15] Youngme Moon. 2000. Intimate Exchanges: Using Computers to Elicit Self-Disclosure from Consumers. *Journal of Consumer Research* 26, 4 (2000), 323–339.

[16] Abhilasha Ravichander and Alan W. Black. 2018. An Empirical Study of Self-Disclosure in Spoken Dialogue Systems. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue* (Melbourne, Australia, 2018-07). Association for Computational Linguistics, 253–263. https://doi.org/10.18653/v1/w18-5030

[17] S. Zahra Razavi, Lenhart K. Schubert, Kimberly van Orden, Mohammad Rafayet Ali, Benjamin Kane, and Ehsan Hoque. 2022. Discourse Behavior of Older Adults Interacting with a Dialogue Agent Competent in Multiple Topics. *ACM Transactions on Interactive Intelligent Systems* 12, 2 (2022), 14:1–14:21. https://doi.org/10.1145/3484510

[18] Anne Roeters. 2018. *Alle Ballen in de Lucht, Tijdsbesteding in Nederland en de Samenhang met Kwaliteit van Leven.* Technical Report 2018-34. Sociaal en Cultureel Planbureau. https://www.scp.nl/binaries/scp/documenten/publicaties/2018/12/07/alle-ballen-in-de-lucht/Alle+ballen+in+de+lucht.pdf

[19] Sherry Ruan, Jacob O. Wobbrock, Kenny Liou, Andrew Ng, and James A. Landay. 2018. Comparing Speech and Keyboard Text Entry for Short Messages in Two Languages on Touchscreen Phones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 159:1–159:23. https://doi.org/10.1145/3161187

[20] Christine Rzepka, Benedikt Berger, and Thomas Hess. 2022. Voice Assistant vs. Chatbot – Examining the Fit Between Conversational Agents' Interaction Modalities and Information Search Tasks. *Information Systems Frontiers* 24, 3 (2022), 839–856. https://doi.org/10.1007/s10796-021-10226-5

[21] Candace L. Sidner, Timothy Bickmore, Bahador Nooraie, Charles Rich, Lazlo Ring, Mahni Shayganfar, and Laura Vardoulakis. 2018. Creating New Technologies for Companionable Agents to Support Isolated Older Adults. *ACM Trans. Interact. Intell. Syst.* 8, 3, Article 17 (jul 2018), 27 pages. https://doi.org/10.1145/3213050

[22] Maarten van Gompel, Axel Vugts, Laurens van der Werff, and Roeland Ordelman. 2016–2022. KaldiNL v0.4.1. https://github.com/opensource-spraakherkenning-nl/Kaldi_NL.

[23] Jelte van Waterschoot, Iris Hendrickx, Arif Khan, Esther Klabbers, Marcel de Korte, Helmer Strik, Catia Cucchiarini, and Mariët Theune. 2020. BLISS: An Agent for Collecting Spoken Dialogue Data about Health and Well-being. In *Proceedings of the 12th Language Resources and Evaluation Conference* (2020-05). European Language Resources Association, Marseille, France, 449–458. https://www.aclweb.org/anthology/2020.lrec-1.57