MOBILE CONVERSATIONAL AGENTS FOR STROKE REHABILITATION THERAPY

ACCEPTED ARTICLE

David Griol Department of Computer Science
Universidad Carlos III de Madrid
Leganés, Spain
dgriol@inf.uc3m.es

Zoraida Callejas
Department of Languages and Computer Systems
Universidad de Granada, CITIC-UGR
Granada, Spain
zoraida@ugr.es

This is a “accepted article” version in the authors employers website, corresponding to the paper:

D. Griol and Z. Callejas, "Mobile Conversational Agents for Stroke Rehabilitation Therapy," 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), Cordoba, Spain, 2019, pp. 513-518, https://doi.org/10.1109/CBMS.2019.00104.

This “accepted article” follows IEEE post-publication policy: https://journals.ieeeauthorcenter.ieee.org/become-an-ieee-journal-author/publishing-ethics/guidelines-and-policies/post-publication-policies/

Abstract

Mobile health (m-Health) has emerged as a rapidly developing area that is transforming clinical research and health care on a global scale. In this paper, we describe a conversational app for the therapy of stroke rehabilitation. The main objective of the conversational app is to help recovering cognitive abilities of patients by means of a set of proposed exercises, which are divided into 8 categories focused on specific abilities. These categories have been defined after a detailed review of the guidelines for rehabilitation and training therapies. In addition, the application integrates a multimodal conversational interface to facilitate human-computer interaction, which has been specially designed for the elderly and patients with motor or visual or disabilities. The exercises provided by the application can be easily adapted to the specific users' requirements and preferences by means of the incorporation, deletion or modification of routines stored into a specific database isolated from the logic of the application.

Keywords Ambient Assisted Living · M-health · Stroke · Conversational interfaces · Mobile devices · Multimodal interaction.

1 Introduction

Stroke is a brain injury that is caused by the reduction or occlusion of blood flow in any area of the brain. More than 120,000 people suffer a stroke in Spain each year. In 2018, it was the first cause of mortality in Spanish women and the second among men[1]. Strokes can occur at any age, and it is one of the ten most frequent causes of children mortality, although the risk increases for people older than 55. Mortality in recent years has been reduced, while the number of people suffering from a stroke has increased and will continue to rise: an increment of 27% has been estimated for the next 25 years, although 90% of cases can be prevented[2]. Most people who suffer from stroke have sequels, either on physical or psychological levels: Motor control (affecting the mobility of the patient, paralysis, involuntary movements, etc.), communication (affecting both verbal and written language), cognition (affecting the ability to learn, reflect, decide and reason), behavior (affecting emotions and personality), and information reception (affecting the patients' five senses).

The type and extent of the sequels after a stroke depend on the areas of the brain damaged and the extent of the damage. There are different ways in which the patient can rehabilitate and recover damaged abilities based on different types of therapies: orientation to reality, reminiscence, language stimulation, stimulation of care processes, etc. The increase in computing performance and mobile connectivity have strengthened the possibilities for mobile health technologies (m-Health).

Conversational interfaces in healthcare provide a more intuitive and natural and intuitive human-machine interaction [1, 2]. In addition, these interfaces allow the combination of spoken, visual, and tactile input and output modalities currently available in widespread devices (e.g., smartphones and tablets) [2]. These factors enrich the participation of patients in their own treatment process, increase and personalize the assistance provided in healthcare services, and improve the patients' outcomes.

In this paper, we present a conversational app for the therapy of stroke rehabilitation in the Android operating system, which main objective is to help stroke patients in their non-pharmacological rehabilitation by stimulating the brain with different kinds of exercises to recover forgotten and damaged skills. The exercises provided by the app have been classified into eight modules, each of them aimed at a group of abilities that may be affected by this disease: language, cognitive, executive functions, orientation, attention, emotional self-control, gnosia, and mobility.

The rest of this paper is organized as follows. Section 2 presents the related work and motivation of our proposal. Section 3 details the multimodal mobile application developed for helping patients who have suffered a stroke. Section 4 shows the experimental process followed to assess the app and the results that have been obtained. Finally, Section 5 summarizes the main conclusions and future research guidelines.

2 State of the art

Figure 1 shows the main types of strokes (ischemic strokes and cerebral hemorrhages). As can be observed, the types of strokes and their sequels can be classified into two large groups: ischemic stroke and cerebral hemorrhages.

[1]Federación Española del Ictus (Last access: May 2019). Available: https://ictusfederacion.es/infoictus/codigo-ictus/

[2]World Stroke Campaign (Last access: May 2019) Available: https://www.worldstrokecampaign.org/

Ischemic stroke is the partial or total obstruction of one or several arteries that carry blood from the heart to the brain. This causes an insufficiency of the blood supply by suspending the flow of oxygen and nutrients necessary to maintain the functioning of the brain cells. They represent between 80-85% of all strokes. Two main types are considered: focal cerebral ischemia and transient ischemic attack. The former affects only a part of the brain when a blood clot has obstructed a brain vessel. The latter is reversible and there is no neurological deficiency when it ends.

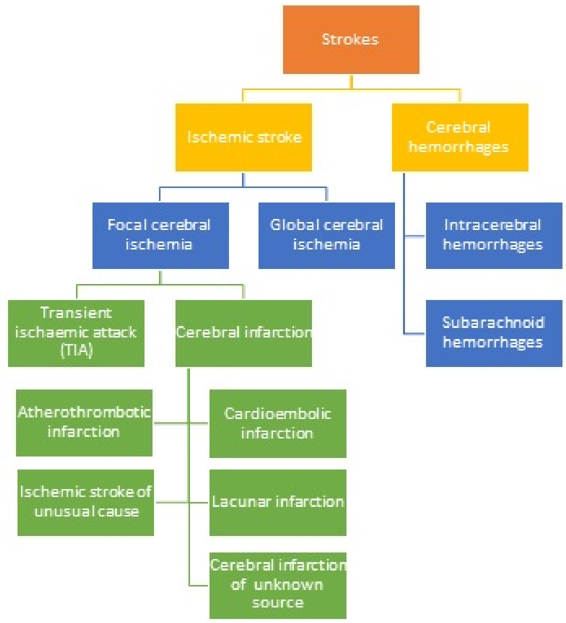

Figure 1: Outline of the main stroke types

After suffering a stroke, in the majority of cases, patients need non-pharmacological rehabilitation, since they need to recover skills that they have forgotten or that have been damaged after the stroke. These skills can be classified into[3]:

- Motor control: sphincter control, paralysis of the limbs, hemiplegia, hemiparesis, dysphagia, spasticity, etc.
- Verbal and/or written communication: aphasia (problems to understand verbal language), alexia (awkwardness to read), anomia (problems to recognize everyday objects), dysarthria (difficulty in articulating sounds and words), dysphonia (problems in the vocal cords that make it difficult to emit a voice).
- Cognition: Cognitive problems can affect learning, reasoning, memory, problem solving, decision making, etc.
- Conduct: Patients who have suffered a stroke may have an emotional imbalance, laughter, sudden crying or depression. In addition, they usually present problems in their daily lives activities that make frustrate them.
- Information reception: It affects sensitive human channels, such as balance, view, hearing, touch and smell.

Patients are taught new ways to perform tasks to compensate the damage caused by stroke. Professionals confirm that for a good rehabilitation, prescribed tasks must repeated to stimulate the brain areas affected. The tasks will be progressively complicated, until the patient has enough autonomy.

There are several non-pharmacological techniques for the recovery of patients. Reality-oriented therapies help patients to achieve a greater interpretation of what surrounds them, increasing self-esteem and feeling of control. It enables the patient to be aware of their spatial, temporary, and personal orientation. Reminiscence therapy help patients to improve their memory, especially by stimulating past events and reviewing their autobiography. Patients are encouraged with questions, photographs or videos of their past, objects that have had a relationship with them, etc.

Language stimulation therapies preserve and increase verbal comprehension. In these therapies, patients need to work especially the attention tasks, which are applied at different levels (basic, sustained attention, selective attention, or divided attention).

[3]National Institute of Neurological Disorders and Stroke (https://www.ninds.nih.gov/) (Last access: May 2019).

The range of possible applications of conversational interfaces in Ambient Assisted Living (AAL) has been considerably extended during the last two decades [3]. Current application domains include the provision of services to change dietary behaviors [4], monitor chronic symptoms [5], provide counseling [6], act as interviewers [7], promote medication adherence [8], detect early potential health disorders or diseases [9], assist on speech therapies [10], etc.

The previously described application domains of healthcare conversational interfaces usually require maintaining a continuous relationship with patients in treatments that may require a long time of care, may involve changes in patients' habits or behaviors, continuous monitoring of chronic diseases, or require helping and assisting them with specific therapies. This continuous relationship implies that dialogs must manage persistent extensive and information about the different patients' interactions with the conversational interface. To enhance this relationship, recent projects have been focused on key aspects of multimodal human-computer interaction, such as memory, proactivity, cognition, social acceptance, memory, continuous learning, affective and emotion recognition [11, 12].

Sánchez-Rodríguez et al. have recently provided a detailed review of mobile applications within the field of neurorehabilitation [13]. Although a total of 69 apps were selected, only 11 apps were specifically designed for neurorehabilitation and just only one was specifically aimed to patients with stroke.

3 The BrainIct App

As described in the previous section, multimodal conversational interfaces require different modules to deal with the recognition, understanding, management, fusion and synthesis of the different input and output modalities. The interaction context is also modeled as key information source that can be unobtrusively acquired using the sensors supported by mobile devices with the Android OS.

In addition to the speech recognition capabilities that are supported by Android devices, it is possible to develop Android applications providing speech interaction (input and output) using the Google Speech API (*android.speech*). Using this API, speech recognition can be carried out using a *RecognizerIntent*, or by means of an instance of *SpeechRecognizer*. The first option initiates the intent and manages the results to complete the recognition, and provides information about the ASR process (e.g., ready to recognize, errors detected, etc). The second option allows a more accurate processing of the speech recognition process by providing developers with a set of recognition events. In both cases, the results are provided by means of an N-best list with confidence scores.

The classes and interfaces required to provide text-to-speech synthesis are provided by the *android.speech.tts* package. This software is used to initialize the TTS engine, control the events related to the the synthesis of a speech utterance, return the synthesized data using the TTS engine, etc.

The dialog interface of the application has been developed by means of the integration of the Dialogflow API[4]. Dialogflow is a Google web service to develop simple conversational agents using natural language interaction. It currently offers support for statistical natural language processing (intents and entities recognition) in 18 languages.

The exercises offered by the application were defined after a detailed review of the main guidelines that describe the set of recommended activities and exercises that should be performed by patients who have suffered a stroke and caregivers to promote the recovery of skills that have been damaged or forgotten. The set of exercises has been divided into 8 categories, each one focused on recovering a specific ability:

- **Language module**: One of the most affected abilities is related to the alteration of oral and/or written language (e.g., difficulties to find the words to continue a conversation, change words by other that sound similarly, problems to repeat words, etc.). In addition, problems with expressive language can be emphasized by troubles to understand them. Figure 2 shows different exercises in this module: to provide a set of words that must start with a given letter (Figure 2(a)), provide a set of words related to a specific description or given condition (Figure 2(b)), read a sentence shown at the screen using also the automatic speech recognizer (Figure 2(c)).
- **Cognitive module**: This category of exercises is focused on recovering the knowledge acquired over time (e.g., memory, intelligence, recognition of nearby people, lack of attention or interest, lack of feelings, etc.). Figure 3 shows different exercises included in this module: to answer several questions about a text (Figure 3(a)), to remember events or daily-life situations, to correctly write a series of words provided by the TTS engine (Figure 3(b)), to select an image that is not related to others provided (Figure 3(c)).
- **Executive functions module**: The decrease in the capacity of the ideomotor, constructive, and/or ideational praxia makes very difficult to carry out simple gestures, use everyday objects appropriately, reduces drawing

[4]https://dialogflow.com/. (Last access: May 2019).

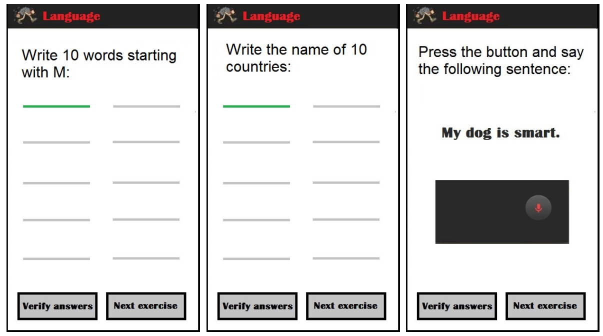

Figure 2: Screenshots of exercises included into the language module

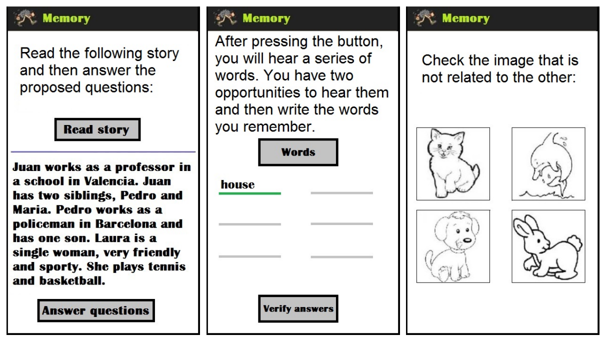

Figure 3: Screenshots of exercises included into the memory module

and writing skills, and affect coordination skills. Figure 4 shows different examples of exercises included in this module: to continue a series with the correct option (Figure 4(a)), to drag coins required to pay a specific amount (Figure 4(b)), or to arrange 6 jumbled images to describe a story (Figure 4(c)).

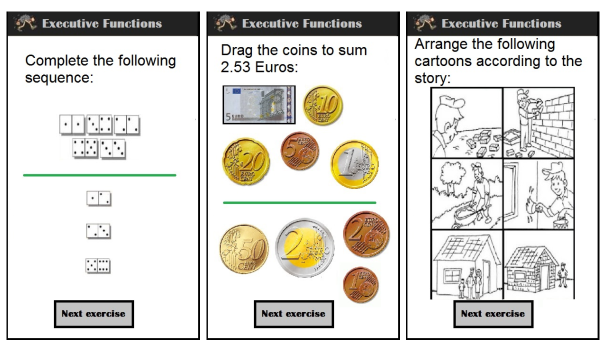

Figure 4: Screenshots of exercises included into the executive functions module

- **Orientation module**: Another main symptom after suffering a stroke is disorientation at three main levels (temporal, spatial, and personal). Figure 5 shows different examples of exercises included in this module: answer questions related to temporal disorientation (Figure 5(a)), select the picture that shows the time in which an activity is carried out usually (Figure 5(b)(c)).
- **Attention module**: This disease also affects the ability to voluntarily refer and apply understanding to a specific objective. Figure 6 shows examples of exercises in this module: to arrange a set of 12 images by dragging them to the correct positions (Figure 6(a)(b)), to select the set of pictures that fulfill the conditions described in the exercise (Figure 6(c)).
- **Emotional Self-Control module**: This disease also affects patients' emotional balance. They may present laughter, sudden crying or depression, changes in their behavior and frustration because they can not solve

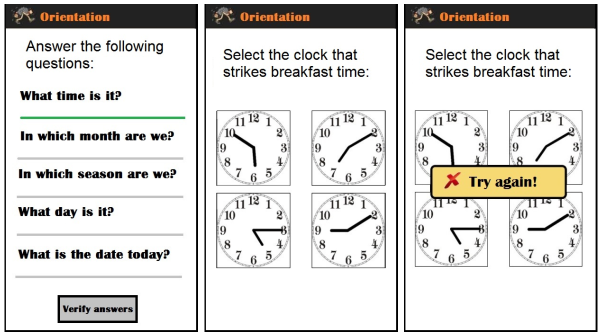

Figure 5: Screenshots of exercises included into the orientation module

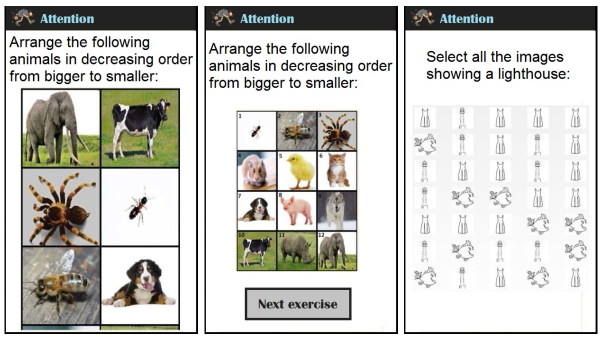

Figure 6: Screenshots of exercises included into the attention module

daily life activities. We have integrated a set of dialogs related to these kinds of activities that are automatically completed and assessed using the dialogflow service.

- **Gnosia module**: Gnosia allows to use our senses to recognize sounds, faces, smells, colors, tastes, shapes, etc. Figure 7 shows different exercises include in this module: to write the color corresponding to daily objects (Figure 7(a)), to recognize the mood of a person (Figure 7(b)), to select the set of pictures included into a specific category (Figure 7(c)).

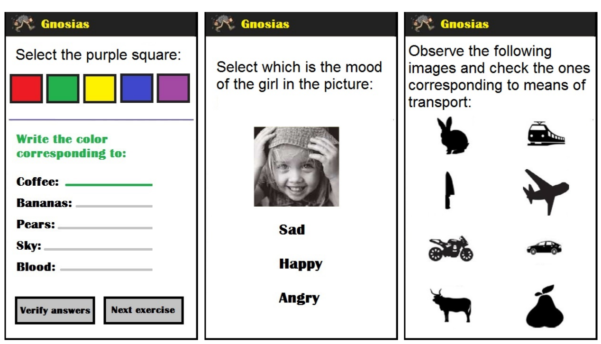

Figure 7: Screenshots of exercises included into the gnosia module

- **Mobility module**: This module is closely related to the previous one. The main objective is to help patients performing exercises that allow them to dress or undress, improve the movement of an arm or a leg, tie shoelaces, walking, etc. We have integrated a collection of videos describing this kind of exercises, which are reproduced using the Youtube Android Player API.

4 Preliminary evaluation

The preliminary assessment that we have completed is focused on the usability, rehabilitative potential and technical quality of the conversational app using the opinions of caregivers and professionals.

4.1 Usability assessment of the app

The usability assessment of the multimodal interaction with the conversational application has been completed using the evaluation questionnaire that is described in Table 1. This questionnaire is based on different standards defined for the evaluation of multimodal interfaces, such as SASSI [14] and AttrakDiff [15]. A Likert scale has been defined to evaluate each one of the 10 questions included in the questionnaire, in which 1 denotes strongly disagreement and 5 denotes strongly agreement. The evaluation was completed by 10 native Spanish users who have not previously suffered a stroke (6 women, 4 men, aged 83 to 54, avg. age 67.3).

Table 1: Questionnaire designed for the usability assessment of the developed conversational app

Previous experience using multimodal interfaces
Q1. Assess on a scale of 1-5 your previous experience using voice interfaces.
Q2. Assess on a scale of 1-5 previous experience using multimodal interfaces.
Understanding of user/system responses
Q3. How well did the system understand your responses?
Q4. How well did you understand the system messages?
Interaction rate
Q5. Was the interaction rate suitable?
Difficulty level using the system
Q6. Indicate the difficulty level of the system.
Presence of errors
Q7. Have you noticed errors during the interaction?
Certainty of what to do at each moment
Q8. Was it easy to decide what to do after each system action?
Rehabilitative objective
Q9. Are the suggested exercises relevant to help stroke patients?
Global satisfaction with the system
Q10. In general, are you satisfied with the system?

Table 2 shows the results of the subjective evaluation using the described questionnaire. We can observe that the participants' previous experience using multimodal interfaces is very varied, as our objective was to evaluate the system with users with different degrees of familiarity with these systems.

The results of the subjective assessment using the previously described questionnaire are shown in Table 2. As can be observed, the users that participated in the evaluation have a varied experience using multimodal interfaces, which makes possible to evaluate the conversational app with different degrees of familiarity with these applications.

Table 2: Results of the usability assessment of the app (For the mean value M: 1=worst, 5=best evaluation)

Question	Avg.	Max.	Min.	Std. dev.
Q1	3.34	4	1	1.28
Q2	3.10	4	1	1.36
Q3	4.94	4	5	0.12
Q4	4.78	4	5	0.31
Q5	4.61	3	5	0.45
Q6	4.73	3	5	0.39
Q7	4.74	4	5	0.47
Q8	4.81	4	5	0.35
Q9	4.77	4	5	0.41
Q10	4.88	4	5	0.32

Regarding the question related to the level of understanding by the system, users evaluate this factor with a very good performance (4.94 out of 5). Users also evaluated that the system responses were understandable (4.78 points out of 5). We use the standard Spanish voice provided by Android.

The interaction rate was also assessed as usually adequate. Several users reported that they would prefer the barge-in mode activated when using speech interaction. Occasionally, the interaction rate was considered slow if a big transaction data was done in the database repositories, but overall, the interaction rate was considered rather fluid (4.61 over 5). Users did not perceive important recognition errors during their interactions with the app (4.74 over 5). With regard the knowledge about the next action to complete during the interaction, the overall satisfaction was 4.81 over 5. Regarding the assessment of the rehabilitative objective, users overall considered that the proposed exercises were oriented to improve the condition of stroke patients and they were adapted to the different damage abilities (4.77 over 5).

4.2 Assessment with professionals and caregivers

We have completed an additional assessment of the conversational interface with the participation of 6 professionals and caregivers. The main aspects that were evaluated included the technical quality, naturalness, and benefits of the app for stroke patients. Table 3 shows the questions included in the questionnaire defined for the evaluation, in which the previously described Likert scale was also used. An additional open question was included to rate the system from 0 to 10 and provide additional remarks.

Table 3: Questionnaire designed for the evaluation of the app with professionals and caregivers

Technical quality
TQ01. The app offers enough interactivity
TQ02. The app is easy to use
TQ03. It is easy to know what to do at each moment
TQ04. The amount of information that is displayed on the screen is adequate
TQ05. The arrangement of information on the screen is logical
TQ06. The app is helpful
TQ07. The app is attractive
TQ08. The app reacts in a consistent way
TQ09. The app provides adequate verbal feedback
TQ10. The app provides adequate non-verbal feedback
TQ11. The feedback provided by the system improves understanding
TQ12. The app encourages continuing using it after errors
Rehabilitative potential
RP01. The app fulfills the objective of helping patients who has suffered a stroke
RP02. The proposed exercises are relevant for this objective
RP03. The design of these exercises was adequate for different kinds of patients
RP04. The activities support significant improvements
RP05. The system promotes a long-term relationship with the users
RP06. The system complements the exercises without distracting or interfering with them

Table 4 shows the results of this evaluation. Users provided a high satisfaction rate with technical aspects. They also perceived the main benefits of the app and relevance of the proposed activities to help patients who have suffered a stroke. The experts also considered the app adequate and attractive. The overall rate for the system was 8.6 (on a scale of 0 to 10).

The open question was used to point out desirable improvements. One suggestion was to change the push-to-talk interface by the possibility of the app constantly listening during the spoken exercises. They also proposed creating a website that would allow users to propose new exercises and different difficulty levels. Similarly, they suggested including new categories of exercises not yet considered in the application. Newly-proposed workouts and exercises should be reviewed by an expert before being incorporated to the application in order to ensure that they are safe to be completed by stroke patients.

Table 4: Results of the evaluation of the application by caregivers. For the mean value M: 1=worst, 5=best evaluation

	Min/max	Average	Std. deviation
TQ01	3/5	4.28	0.56
TQ02	4/5	4.67	0.44
TQ03	4/5	4.82	0.31
TQ04	3/5	4.23	0.62
TQ05	4/5	4.71	0.48
TQ06	4/5	4.85	0.31
TQ07	4/5	4.31	0.52
TQ08	4/5	4.42	0.47
TQ09	4/5	4.83	0.39
TQ10	4/5	4.67	0.52
TQ11	3/5	4.03	0.66
TQ12	3/5	4.11	0.54
RP01	5/5	5.00	0.00
RP02	4/5	4.87	0.16
RP03	4/5	4.33	0.41
RP04	4/5	4.31	0.43
RP05	4/5	4.62	0.37
RP06	4/5	4.74	0.26

5 Conclusions and future work

We have described a mobile conversational interface designed to help patients recovering after a stroke. The multimodal interface provided by the app allows the combination of several input and output interaction modalities that are very useful for older people and patients with motor or visual disabilities. The set of proposed exercises currently provided by the app has been evaluated as relevant for post-stroke rehabilitation in the preliminary evaluation completed by older adults, professionals and caregivers.

As future work, we want to conduct a more detailed assessment of the app with patients who have suffered a stroke. We also want to extend the the app with the ability to recognize and act upon the users' emotional states.

Acknowledgments

This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823907 (MENHIR project) and from the Spanish Ministry of Economy, Industry and Competitiveness through the CAVIAR (MINECO, TEC2017-84593-C2-1-R) project.

References

[1] T. Bickmore and T. Giorgino. Health dialog systems for patients and consumers. *Journal of Biomedical Informatics*, 39(5):556–571, 2006.

[2] Michael F. McTear, Zoraida Callejas, and David Griol. *The Conversational Interface: Talking to Smart Devices*. Springer, 2016.

[3] L. Laranjo, A.G. Dunn, H.L. Tong, A.B. Kocaballi, J. Chen, R. Bashir, D. Surian, B. Gallego, F. Magrabi, A.Y.S. Lau, and E. Coiera. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, 25(9):1248–1258, 2018.

[4] H.K. Delichatsios, R.H. Friedman, K. Glanz, S. Tennstedt, C. Smigelski, and B.M. Pinto. Randomized trial of a talking computer to improve adults eating habits. *American Journal of Health Promotion*, 15:215–224, 2000.

[5] J. P. Migneault, R. Farzanfar, J.A. Wright, and R.H. Friedman. How to write health dialog for a talking computer. *Journal of Biomedical Informatics*, 39(5):276–288, 2006.

[6] Eva Hudlicka. Virtual training and coaching of health behavior: Example from mindfulness meditation training. *Patient Education and Counseling*, 92(2):160–166, 2013.

[7] L. Pfeifer and T. Bickmore. Designing Embodied Conversational Agents to Conduct Longitudinal Health Interviews. In *Proc. IVA'10*, pages 4698–4703, 2010.

[8] T.W. Bickmore, K. Puskar, E.A. Schlenk, L.M. Pfeifer, and S.M. Sereika. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers*, 22:276–288, 2010.

[9] G. Villarrubia, J.F. de Paz, J.M. Corchado, and J. Bajo. EKG Intelligent Mobile System for Home Users. In *Proc. IBERAMIA'14*, pages 767–778, 2014.

[10] O. Saz, S.-C. Yin, E. Lleida, R. Rose, C. Vaquero, and W.-R. Rodríguez. Tools and Technologies for Computer-Aided Speech and Language Therapy. *Speech Communication*, 51(10):948–967, 2009.

[11] S. Payr. Closing and closure in human-companion interactions: Analyzing video data from a field study. In *Proc. IEEE RO-MAN'10*, pages 476–481, 2010.

[12] I. Leite, A. Pereira, G. Castellano, S. Mascarenhas, C. Martinho, and A. Paiva. Modelling empathy in social robotic companions. *Advances in User Modeling*, 7138:135–147, 2012.

[13] M. T. Sánchez Rodríguez, S. Collado-Vázquez, P. Martín-Casas, and R. Cano de la Cuerda. Neurorehabilitation and apps: A systematic review of mobile applications. *Neurology*, 33(5):313–326, 2018.

[14] Kate Hone. Usability measurement for speech systems : SASSI revisited. Technical report, Brunel University, London, UK, 2014.

[15] M. Hassenzahl, M. Burmester, and F. Koller. *Mensch & Computer 2003. Interaktion in Bewegung*, chapter AttrakDiff: A questionnaire for measuring perceived hedonic and pragmatic quality, pages 187–196. Vieweg+Teubner Verlag, 2003.