Griol, D, Molina, J M, Sanchis, A. A multimodal conversational coach for active ageing based on sentient computing and m-health. *Expert Systems*. 2020; 37:e12454.

DOI: 10.1111/exsy.12454

A multimodal conversational coach for active ageing based on sentient computing and m-health

David Griol*, Jose Manuel Molina, Araceli Sanchis

Dept. of Computer Science, Universidad Carlos III de Madrid, Avda. de la Universidad, 30, 28911. Leganés, Spain.

Correspondence
*David Griol, Universidad Carlos III de Madrid.
Email: dgriol@inf.uc3m.es

Present Address
Avda. de la Universidad, 30, 28911. Leganés (Spain)

Summary

As life expectancy increases it has become more necessary to find ways to support healthy ageing. A number of active ageing initiatives are being developed nowadays to foster healthy habits in the population. This paper presents our contribution to these initiatives in the form of a multimodal conversational coach that acts as a coach for physical activities. The agent can be developed as an Android app running on smartphones and coupled with cheap widely available sport sensors in order to provide meaningful coaching. It can be employed to prepare exercise sessions, provide feedback during the sessions, and to discuss the results after the exercise. It incorporates an affective component that informs dynamic user models to produce adaptive interaction strategies.

KEYWORDS:
Sentient computing, active ageing, m-health, conversational interfaces, mobile interfaces, multimodal interaction, affective computing.

1 | INTRODUCTION

The World Health Organization (WHO) defines active ageing as "the process of optimizing opportunities for health, participation and security to enhance quality of life as people age allowing them to realize their potential for physical, social and mental well-being throughout the life course" (WHO 2002). According to this organization, the proportion of people over 65 years will double between 2010 and 2050, and the age group that is expected to grow the most is those over 80 (WHO 2018). These changes are fundamentally changing population demographics internationally, resulting in increasingly ageing populations. Thus, it is key issue to ensure that this increment in life expectancy, due to advances in medicine and healthcare, is coupled with a good quality of life (van Dyk, Denninger, & Richter 2014).

The concept of active ageing is concerned with extending healthy life expectancy in several different ways (Lionis & Midlov 2017; Martínez-Tomás, Fernández-Caballero, & Ferrández 2014; Robbins, Keung, & Arvanitis 2018). The WHO in a report to define a global strategy and action plan on ageing and health from 2016 to 2020[1] indicated different guidelines to make active ageing possible, the first being the promotion of "good health and healthy behaviors at all ages to prevent or delay the development of chronic disease". These include being physically active, eating a healthy diet, and avoiding alcohol and tobacco. Also they highlighted the need that the strategies developed should take into account the constraints of low- and middle-income health systems.

We are particularly interested in the guideline of promoting physical activity (Foster, Richards, Thorogood, & Hillsdon 2013; Muellmann, Forberger, Mollers, Zeeb, & Pischke 2018). Currently there are different technologies being developed to help users achieve this goal. For example, there exist many commercial applications, especially for runners, which rely on sensors that offer feedback about the user's achievements after the exercise sessions. Most of these applications are rather static in the sense that they offer the information after the sessions, mainly in the form of statistics and graphics. Also, they follow a traditional human-computer interaction (HCI) approach in which the interaction with the user is limited.

[0]**Abbreviations:** AAL, Ambient Assisted Living; ASR, Automatic Speech Recognition; ATP, Adenosine TriPhosphate; BLE, Bluetooth Low Energy; GUI, Graphical User Interface; m-health, Mobile health; TTS, Text-to-Speech; WHO, World Health Organization.

[1]http://www.who.int/ageing/WHO-GSAP-2017.pdf?ua=1. [Accessed December 2018].

As people age, their health needs tend to become more chronic and complex. Thus, a transformation is needed in the way that health systems are designed to ensure affordable access to integrated services that are centered on the needs and rights of older people. Digital strategies, such as e-health, m-health and telemedicine, offer the potential to deliver active ageing in a cost-effective manner at scale (Robbins et al. 2018). As will be discussed in Section 2, research systems offer more natural interaction through the use of conversational interfaces (G. Lee, Kim, Jeong, & Kim 2015; McTear, Callejas, & Griol 2016; Pieraccini 2012) that allow speech-based interaction. However, in this case there is a lack of dynamic feedback while the exercise is being performed, as most agents do not incorporate sentient computing.

The proposal in this paper is an architecture to build conversational coaches to encourage healthy habits, monitoring the user's health during the exercises, planning appropriate exercise routines, and supporting users to achieve their goals. The system employs easily accessible and widely available sensors that can be used together with mobile speech technologies in order to foster and control regular exercising at an appropriate pace for each user. Our proposal adapts the interaction metaphor to the different conditions in which the system can be used, and includes an affective component that allows the user provide feedback about how they felt during the sessions, which helps to build a complex user model for a more intelligent system behavior. Thus, unlike current systems that work with biosignals, it is not about presenting the measured data directly to the user, but about maintaining dialogues at a higher level of abstraction directly about the observed behaviors, the predicted behaviors and their possible consequences.

The proposed architecture has been used to develop a m-health multimodal virtual coach for Android smartphones that fosters active ageing among elderly people. This mobile application allows users to perform different types of workouts, depending on the diseases they suffer, if any. To this end, spoken interaction and emotion recognition have been incorporated in the main modules to provide the user with an enhanced multimodal interface that also monitors and processes biosignals by means of bluetooth-enabled devices.

The remainder of the paper is as follows. Section 2 describes the related work regarding the potential of digital technologies to enable active aging and, specially, promote physical activity in the context of Ambient Assisted Living (AAL). This section also describes the main concepts related to the definition of training programs for the elderly and the use of conversational interfaces to provide AAL services. Section 3 describes the proposed framework to develop adaptive multimodal AAL interfaces and its application to develop a mobile multimodal virtual coach to promote physical exercise among elderly people. Section 4 describes the interaction modes with the coach application. Section 5 presents the experimental set up and the results of the evaluation of the virtual coach. Concluding remarks and directions for future work follow in Section 6.

2 | RELATED WORK

Related to the increasingly ageing populations described in the previous section, there is an increase in the use of digital technology by the elderly. In fact, older adults now represent the fastest growing population of adopters of Internet and computer technologies (Chiu & Liu 2017). Therefore, there is a clear potential to adapt digital technologies, such as telemedicine and telehealth, to the challenge of enabling active aging (Christophorou et al. 2016; Dimitrova 2013; Richard et al. 2016).

In addition to applications that target physical activity, the kinds of applications designed for active ageing cover different types of interventions: information systems for clinical use (Bourret & Bousquet 2013), coaching of the elderly person (Jongstra et al. 2017; Richard et al. 2016; Tiedemann et al. 2016), cognitive behavior therapy (Preschl, Wagner, Forstmeier, & Maercker 2011), cognitive training (Gates et al. 2016; Reijnders, Geusgens, Ponds, & van Boxtel 2017; Santos, Reis, & Barroso 2013), communication tools between healthcare professionals and patients or carers (Henriquez-Camacho, Losa, Miranda, & Cheyne 2014; Preschl et al. 2011), decision support systems (Bourret & Bousquet 2013), education of patients (Henriquez-Camacho et al. 2014; Lattanzio et al. 2014), applications that target ethical considerations of technologies (Karki, Savel, Sallinen, & Kuusinen 2013), computer and serious games (Gates et al. 2016; Santos et al. 2013), information systems and technology development (Bourret & Bousquet 2013; Ferreira, Sayago, & Blat 2017; Karki et al. 2013), applications focused on the management and coordination of e-health interventions (Lattanzio et al. 2014), applications targeting rehabilitation (Kaufman 2012), systems that target social inclusion and participation (Ferreira et al. 2017; Henriquez-Camacho et al. 2014), remote care technologies (Jongstra et al. 2017; Knowles et al. 2016), telehealth applications to help support self-management (Henriquez-Camacho et al. 2014; E. Lee, Han, & Jo 2017; Tiedemann et al. 2016), and telemedicine systems for direct intervention of the clinician (Keijser et al. 2016; Lattanzio et al. 2014).

Traditionally, these kinds of applications have been designed for in-home use in the context of assisted living, where users can, for example, communicate some vital signs to a system that logs them, provides feedback and/or forwards them to a doctor (Gronvall & Verdezoto 2013), or obtain advice (Hudlicka 2013). However, advances in the development of smartphones and their spread among the population provide interesting opportunities to migrate such functionality to m-health apps that offer the possibility to make more meaningful and dynamic interventions (Muellmann et al. 2018; Poole 2013).

2.1 | Physical activity in the old age and exercising applications

Regular physical exercise has a positive impact in the lives of old people, including the following benefits (Cvecka et al. 2015): improve the ability to take care of oneself; lead to general well-being; keep the senses agile; introduce positive lifestyle changes; reduce anxiety, insomnia, and depression; increment aerobic capacity, muscular strength and flexibility; decrease the risk of a cardiovascular disease.

Physical condition is essential when it comes to prescribing physical exercise. The physical condition depends on the muscular strength, the aerobic endurance, the flexibility and the balance (Cvecka et al. 2015; Muellmann et al. 2018). Therefore, a thorough evaluation of the physical condition is essential to determine the adequate exercise for an old person.

The most relevant aspects to evaluate are cardiorespiratory endurance, body composition, strength and muscular endurance, flexibility, balance and coordination. In this respect, it can be seen that elderly people integrates a very diverse group of people depending on their functional capacities, that is, their autonomy. This consideration allows to classify elderly people according to three profiles: physically independent, in a special situation of fragility, and physically dependent (require caregivers).

Two main types of physical activity are usually considered (Smith 2005). Aerobic (dynamic) activity involves exercises of low or medium intensity and long duration. It has been demonstrated that aerobic activity has a positive impact at cardiovascular level, helping to prevent the development of cardiovascular diseases. More specifically, both the heart rate and the blood pressure tend to decrease over time (Pinto 2007). Anaerobic (static) activity involves exercises of high intensity and short duration. The body does not require oxygen because it obtains energy from immediate sources, such as ATP or glucose. It involves specific muscle groups, producing changes in the muscular tension but without changing the length of the muscle fiber significantly (Fleg 2012). It has been proven to improve strength and muscular endurance, being appropriate for body toning. The recommended physical activity for elderly people is aerobic exercise of low intensity and that has the least impact possible on joints (Fleg 2012). This includes exercises such as walking, marching, pedaling, swimming, or rowing.

A training program for the elderly must be balanced and include activities to achieve each of the three goals of a good physical condition: increase flexibility, increase strength and increase cardiovascular endurance. Therefore, the main goal of a program is to keep the general functional capacity and integrity of the locomotor system as well as the prevention, treatment and rehabilitation of certain diseases (Golbidi & Laher 2012). In this way, an old person can remain independent for a longer period of time and with a better functional capacity. Moreover, doing regular exercise can delay the decay of the nervous system, which will enhance in turn the agility and the reflexes.

A workout can be defined by the following four components (Bhatia & Sood 2017):

- Intensity: degree of effort required by the exercise. It can be calculated objectively by means of a stress test or subjectively using the Börg scale, which measures the effort feeling of the person doing the exercise.
- Duration: amount of time doing exercise. It is recommended to have sessions from 20 to 60 minutes.
- Frequency: number of times a week that the exercise is done. It is recommended to have sessions from 3 to 7 times a week.
- Type of activity: kind of exercise to be done during the session.

The prescription of a workout includes information about the three components of a session, namely the warm-up, the exercise, and the cool-down, and it is designed to adjust to the state, health requirements and medical problems of a patient. Every session is advised to start with a warm-up period of at least 5 minutes to prepare the body progressively for a higher intensity exercise, thus preventing injuries in the locomotor system. Warm-ups usually include stretching exercises for the muscles. Finally, the exercise should finish with a 5-minute cool-down to go back to calm.

The sessions must also be carefully planned to adjust to the physical condition of the elderly. In this regard, the minimum duration advised is around 20 minutes to ensure that the aerobic system reaches its full-working mode. Then, between 1 and 3 minutes can be added every week. Also, the frequency prescribed should be at least twice a week and up to four times a week (always in alternating days).

In the systematic review recently presented in (Muellmann et al. 2018), e-Health interventions were found effective to promote physical activity in older adults aged 55 years and above when compared to no intervention control groups, at least in the short term. Regarding the intervention intensity necessary for behavior change, the main conclusions indicate that greater engagement with web and mobile-based interventions is associated with larger effects on physical activity.

Sophisticated industrial sport coaches (e.g. by Adidas™, Garmin™, or Nike™) include personalized workouts and performance tracking. Three of the most well-known exercise applications for mobile devices are Endomondo[2], Runtastic[3] and Freeletics[4]. Each of them employs a different

[2]https://www.endomondo.com/. [Accessed December 2018].
[3]https://www.runtastic.com/. [Accessed December 2018].
[4]https://www.freeletics.com. [Accessed December 2018].

approach to training. However, these applications are not aimed at older populations, and the interaction modes are restricted. While exercising, some of such applications provide voice messages, but there are no conversational features and the user's affect is not considered in the design of the training plans.

2.2 | Conversational interfaces and E-health

Conversational interfaces can be defined as automatic systems that emulate a human being in a dialog with another person, in order to complete a specific task (Griol, Callejas, López-Cózar, & Riccardi 2014; McTear et al. 2016; Pieraccini 2012). During the last two decades, these interfaces have been increasingly used in AAL to provide a more natural and intuitive human-machine communication, promoting patients' involvement in their own care, assisting in health care delivery, and improving patient outcome (T. Bickmore & Giorgino 2006; Laranjo et al. 2018). Typical application domains include providing services such as interviews (Pfeifer & Bickmore 2010), counseling (Hudlicka 2013), chronic symptoms monitoring (Giorgino et al. 2004; Mooney, Beck, Dudley, Farzanfar, & Friedman 2004), medication adherence (Allen et al. 2006; T. Bickmore, Puskar, Schlenk, Pfeifer, & Sereika 2010), and the adoption of healthy behaviors (Farzanfar, Frishkopf, Migneault, & Friedman 2005).

Multimodal conversational interfaces go an additional step beyond traditional graphical user interfaces (GUIs) by adding the possibility to communicate with widespread devices, such as smartphones and tablets through other interaction modes, such as speech, tactile, and visual interaction. However, we often find a dichotomy between mobile and home applications which rarely combine to provide more dynamic services in different contexts. Also we can differentiate between applications that are focused on monitoring, for example blood pressure for diabetes (Harper et al. 2008), and others that are more focused on counseling (T. W. Bickmore, Schulman, & Sidner 2013 2011), or planning, e.g. plan and remind medication intake (Dalgaard, Gronvall, & Verdezoto 2013).

The application of these interfaces in the healthcare domain implies maintaining a continuous relation with patients through time with the main aim of either eliciting changes in patients' behaviors or habits, monitoring chronic-conditions or assisting them under a determined therapy. This continuity forces dialogs to manage extensive and persistent information about the different sessions of the patient. In addition, recent projects have also addressed very important aspects in multimodal human-machine interaction like the social acceptability of verbally assistive systems (Payr 2010) or the possibility of including additional capabilities such as memory, cognition, emotion recognition or learning (Cavazza, de la Cámara, & Turunen 2010; Leite et al. 2012; Young 2011), Soprano (Sixsmith et al. 2009), Humaine (Andre et al. 2011; Rehrl et al. 2013)).

In this paper, we propose a single framework to develop a multimodal conversational coach that operates in two main modes (exercise and conversation), combine the benefits of mobile and in home applications, and fulfill the objectives of monitoring, planning and coaching. The first interaction mode is active while the user is engaged in exercise, for example, going for a walk. In this mode the system monitors and provides multimodal feedback to the user, while it updates the user model with their achievements. The second mode corresponds to a more relaxed situation in which the user can talk to the agent to plan new activities or discuss their progress; an affective component makes it possible for the system to adapt dynamically its strategy considering the users' feedback.

3 | THE PROPOSED MULTIMODAL CONVERSATIONAL COACH

Figure 1 shows the architecture proposed to develop a multimodal conversational coach to promote physical activities in active ageing. As it can be observed, different modules, sensors and models cooperate to prepare exercise sessions, include conversational interaction and affective computing, provide feedback during the sessions, and to discuss the results after the exercise.

In the first stage the *user state composer* takes the biosignals and the user model as input and constructs the user state which is forwarded to the *advice manager*. This manager is a specialized dialog manager that does not require any user input and that computes the best coaching advice taking into account the the user and training models. For example, walking for 40 minutes at a high speed may be good for some users, but too exhausting for others. Thus, the advice depends on the type of exercise, the dynamic user state, and the user characteristics.

Finally, once the output is decided, the *multimodal output generator* generates an appropriate feedback to the user which is visual (through graphics and texts in the smartphone screen), spoken (through synthesized speech) and tactile (through vibrations). The user is able to configure the multimodal output so that it adapts to their preferences and needs.

The following subsections describe the sensors and mobile technology required for the implementation of the multimodal coach, the models defined, and the guidelines followed for the workouts design. Section 4 describes how the interaction is controlled and the main operation modes of the coach.

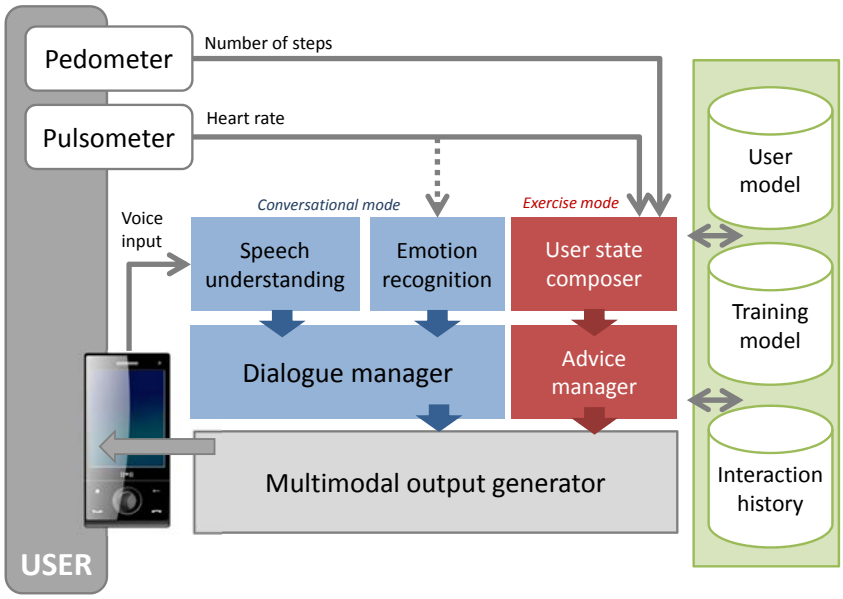

FIGURE 1 Architecture of the multimodal conversational coach

3.1 | Speech recognition and synthesis

Besides the recognition capabilities that are implemented within the Android operating systems, there is the possibility to build Android apps with speech input and output using the Google Speech API (package *android.speech*). With this API, speech recognition can be carried out by means of a *RecognizerIntent*, or by creating an instance of *SpeechRecognizer*. The former starts the intent and processes its results to complete the recognition, providing feedback to inform that the ASR is ready or there were errors during the recognition process. The latter provides developers with different recognition related events, thus allowing a more fine-grained processing of the speech recognition process. In both cases, the results are presented in the form of an N-best list with confidence scores.

ASR is available via *android.speech.SpeechRecognizer*. The way the speech recognition works is as follows:

- The speech recognition starts when the listen method is called. The listen method creates a *RecognitionListener* to listen to the user, specifying the language model and the maximum number of results. Finally, the intent is started using the *SpeechRecognizer* object.
- When the recognition finishes, the *onResults* method is automatically called (it is overwritten from *RecognitionListener*). The *onResults* method receives the results of the recognition and forwards them to the *processArsResults* abstract method. *I*n case of errors during the recognition process, the *onError* method is automatically called (it is overwritten from *RecognitionListener*). Again, the *onError* method receives the error code and forwards it to the *processAsrError* abstract method, so that each class extending Speech can handle the error in the most appropriate way.

TTS is available via *android.speech.tts.TextToSpeech*. The way the speech synthesis works is as follows:

- The *TextToSpeech* object is ready to use when the *onInit* method is called. When the synthesizer is ready, the language of the *TextToSpeech* object is set to English (UK) and the abstract method *onTTSReady* is called.
- The speak method can be called to ask the *TextToSpeech* object to utter a message for which an utterance ID has been assigned. The utterance ID is required to track the message and get signaled by the *onDone* method.
- The *onDone* method calls the abstract method *processTtsResults* so that any class extending Speech can decide how to handle the end of an utterance.

3.2 | Sensors and mobile technology

The sensors employed in our architecture belong to two types. The first type involves sensors that are already incorporated in the smartphone running the coach. These include the microphone and GPS. The second type involves sensors that are external to the smartphone, such as the pulsometer and pedometer. These sensors measure biosignals: pulsometers offer information about heart rate and skin conductivity, whereas pedometers count the number of steps during walking. They are widely available, in the form of easily wearable devices, such as watches and bands. One of the key points of our proposal is that the user can choose the brand and type of sensors they prefer, so it is not necessary to buy any special or expensive equipment for the coach application.

Bluetooth Low Energy (BLE) is a wireless personal area network technology developed by the Bluetooth Special Interest Group[5] in the advent of the Internet of Thing (IoT). Its purpose is to provide the same functionality and communication range as classic Bluetooth, but with a much lower energy consumption and cost, which is ideal for devices that have to run for long period of time on small power sources. The most important mobile operating systems, such as Android, iOS and Windows Phone, natively support Bluetooth Smart. Similarly, a wide range of mobile devices currently support this technology.

3.3 | User, training and history of interaction models

As Figure 1 shows, in our proposed architecture there are three main knowledge models: the user model, training model, and interaction model. The user model contains "static" information (name, age, and gender), as well as information that is updated according to the exercise sessions and interactions. This dynamic part deals mainly with the user's progress in the exercises and the user's affect in relation to this progress. To measure the user progress we have considered the frequency of use of the coach application, the frequency of exercise sessions, and the extent to which the objective of each session is achieved (e.g. if the user had to cover a certain distance, what is the percentage covered), heart rate register, distance register, time register, and rhythm register.

The training model codifies the best strategies for the coach according to the physical condition of the user, and the expected behavior of the user in positive and negative scenarios. This knowledge is provided by expert trainers and codified as rules that are reinforced with suggestions from the scientific literature (e.g. from Journal of the American Geriatric Society, Journal of Aging and Physical Activity and Journal of Exercise Science and Fitness). With respect to the affective model, we have considered a simple model of polarity and arousal (positive vs. negative affect and intensity level specified as a percentage), which is coupled with the information of the exercise session to which it relates. This information is codified as tuples *(session id*, *expected results*, *observed results*, *expected affect*, *observed affect)*.

The history of the interaction is saved as logs including the date, interaction mode (see Section 4), log of sensor information, and log of sensor communication errors. Moreover, when there is a conversation with the agent we trace interaction parameters (duration of the dialog, number of user turns, number of system turns, information about the speech recognizer performance, speech understanding success rates), and emotion recognition results.

3.4 | Workouts design

In the bibliography, several different workouts can be found that are connected to specific diseases suffered by elderly people. We decided to focus on independent users (as opposed to dependent, fragile users who required external care) and in aerobic activities according to the prescription recommended by the WHO[6]. Our proposal focuses on using physical exercise as a mean to improve physical condition rehabilitating the person from a particular disease. Table 1 compiles the list of diseases considered.

As it has previously described, a workout is characterized by its intensity, its duration, its frequency, and the exercise it consists of. To provide users with a flexible range of workouts, they can select a specific exercise based on its intensity. The exercises considered in the application includes walking (low impact), cycling (low impact), skating (low impact), dancing (medium impact), and running (medium impact).

Considering that the workouts are aimed at the elderly population, and especially to people with some sort of ailment, we have not considered high impact exercises. Moreover, since the heart rate of the user will be measured to account for the user's performance, some exercises have not been considered because they are not compatible with wearing a Bluetooth monitor, such as swimming. Based on the results exposed by the WHO, the workouts in Table 2 have been devised.

[5]https://www.bluetooth.com/. [Accessed December 2018].
[6]https://www.who.int/dietphysicalactivity/factsheet_olderadults/en/. [Accessed December 2018].

TABLE 1 Diseases considered in the workouts

Disease	Type
acute myocardial infarction (a.k.a. heart attack)	cardiovascular system
heart failure	cardiovascular system
hypertension (a.k.a. high blood pressure)	cardiovascular system chronic obstructive
pulmonary disease (mild)	respiratory system
chronic obstructive pulmonary disease (severe)	respiratory system
diabetes mellitus	metabolic
arthritis - osteoarthritis	systemic
fibromyalgia	systemic
osteoporosis	locomotor system

TABLE 2 Workouts included in the multimodal conversational coach

DISEASE	EXERCISE INTENSITY	INTENSITY	DURATION	FREQUENCY
Acute Myocardial Infarction	medium	40-60% HRR (Börg: 11-13)	30-60min	2-3 sessions/week
Heart failure	low	Börg: 14→Börg: 10	10min→5min	1-2 sessions/week
Hypertension	medium	40-70% HRR (Börg: 11-14)	30-60min	4-5 sessions/week
Chronic Obstructive Pulmonary Disease (severe)	low	40% HRR (Börg: 17)	3 series of 10min	3-5 sessions/week
Chronic Obstructive Pulmonary Disease (mild)	low	40-65% HRR (Börg: 17)	20-40min	3-5 sessions/week
Diabetes mellitus	low	40-60% HRR (Börg: 12-13) with changes to 70% HRR (Börg: 15-16)	20-30min	Daily sessions
Arthritis - Osteoarthritis	low	40-75% HRR (Börg: 12-16)	20-60min	3-5 sessions/week
Fibromyalgia	low	30-50% HRR (Börg: 7-12)	20-40min	2 sessions/week
Osteoporosis	medium	40-60% HRR	20-60min	3-5 sessions/week

There are two main types of workouts: continuous and intervallic. Continuous workouts consist in doing an exercise for a given duration, whereas intervallic workouts consist in doing several series of exercises, each series for a given duration, repeating the sequence of series a number of times. Thus, both types of workouts have to be reconciled so that they can be modeled following the same pattern. The strategy we conceived was the following: each workout, no matter which type, will be defined by the frequency and the exercise perform and it will consist of a number of series repeated a number of times. In this way:

- Continuous workouts consist of a unique series, with one repetition.
- Intervallic workouts consist of several series, with one or more repetitions.

Using this model it is possible to describe a more dynamic range of exercises than just using the continuous exercise model.

4 | COACH OPERATION MODES

The coach application consists of different modules. The login/register module (see Figure 2) allows users to register and log in. The register module allows users to fill in a registration form, providing his personal information. The profile module displays the user's personal information and allows to edit the profile and change the password. The start module acts as the main menu of the application, providing access to all its functionalities, such as viewing the profile, logging out, training or managing the current workouts.

The multimodal coach has two main operation modes: *exercise* and *conversation*. The first mode is used during exercise, and the second before and after the exercise in order to prepare and discuss the sessions.

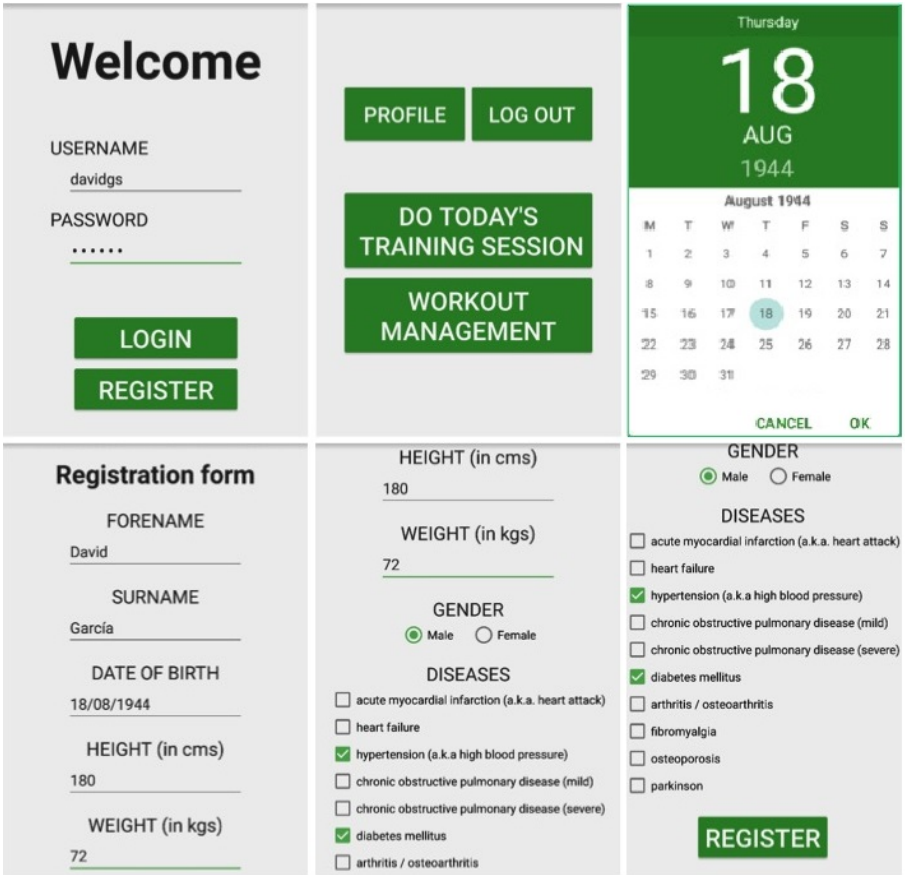

FIGURE 2 Login / Register, start and DOB date picker screens

4.1 | *Exercise* mode

The first mode corresponds to the training sessions. During exercise, the user wears the sensors, which communicate with the smartphone via bluetooth. The system collects the biosignals at run time and updates the user state. When a significant change occurs either because there is an important change in the user state, or because they have reached a target in the training session, the coach provides feedback to the user.

Systems that provide feedback based on biosignals (e.g., smart watches for runners) usually present just statistics and graphics, and often only when the training session is finished and not during it. In our proposal, the coach is able to provide feedback to the user which, on the one hand, is tailored to the user's needs whilst on the other provides reinforcement for the exercise, giving advice and support to the user (e.g., how to continue or when to stop).

The training module allows users to complete the workout session scheduled for the day. This functionality requires connecting BLE devices (e.g., the heart rate monitor to measure their pulse). In addition, the coach informs users about their progress and encourages them with spoken messages.

The user interface shows the following information pieces and control options (see Figure 3): current series and the current repetition, time left for the current series of the session to end, evolution of the current series of the session, a chronometer to count the time that the user takes a break, current heart rate of the user, manually scan for BLE devices in case none is discovered automatically, information about the discovered BLE devices, start the session (or resume it after a break), pause the session and take a break, and finish the session before the time is up.

The functionality of this mode is the most complete of the coach application, since it involves many different actions: speech recognition and synthesis, dialog management, BLE device discovery and connection, access to database repositories, etc.

First of all, when the Training activity is launched, the user's prescribed workouts are retrieved. If the user does not have any currently active workout, they are informed about this and the application returns to the Start menu, since there is no workout to be done. Otherwise, the user's prescribed workout series are retrieved to complete the training session. Then, the application scans for BLE devices. Similarly to what happens when the user launches the application after having previously registered or logged in, the UUID of the heart rate monitor is recorded in the application's shared preferences after its first use, so that it will be automatically recognized the next time.

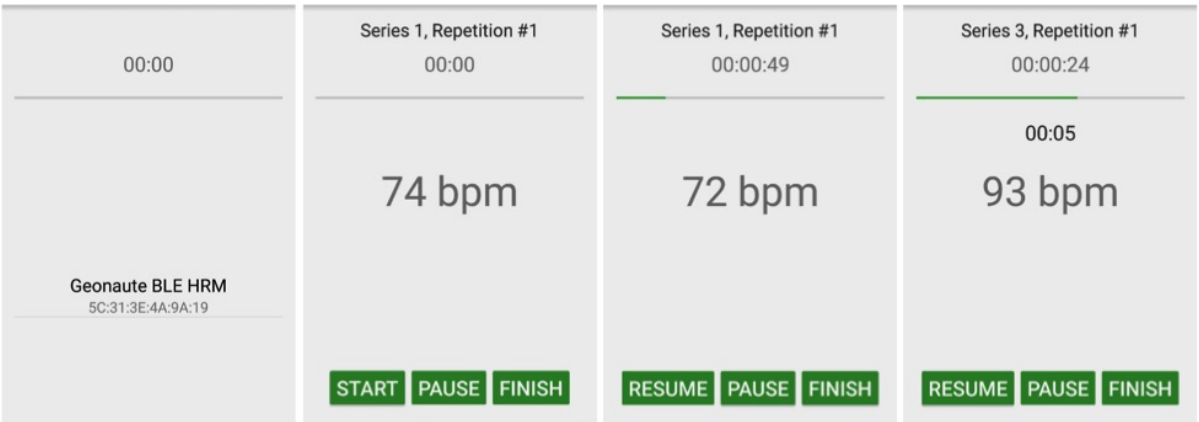

FIGURE 3 Automatic scan of BLE devices(1), training screen ready for the workout (2), while the workout is being done (3), and when the user is taking a break (4)

After that, the user's resting heart rate is retrieved, as it will be required to determine the heart rate reserve later on. The resting heart rate is measured by averaging the heart rate of the user measured during one minute, while the user stays still. Users can control the training session using the Start, Pause and Finish buttons.

Once the session starts, the progress of the session is measured with a *CountDownTimer* and a *ProgressBar*, which display the progress both numerically and graphically. Moreover, when the user pauses the session to take a break, this time is measured with a chronometer.

At the beginning of each series of the session, the heart rate reserve associated to that series is computed, since the heart rate reserve depends on the intensity of the exercise. The heart rate reserve can be computed with the following formula:

$$HRR=(\text{maximum HR-resting HR})*\text{intensity}+\text{resting HR}$$

Therefore, the following parameters are required:

- Maximum HR: it can be derived from the user's age Maximum HR=206.9-0.67*age (Gellish et al. 2007).
- Resting HR: it is measured the first time the user starts a training. The user is asked to stay still and relaxed for one minute and the system records their pulse during this time. The resting HR is calculated as the average of those values.

The oral interaction in this module is mainly one-way: the application talks to the user to provide them information while they train (so that they do not have to look at the screen). However, when a series session is over, the user is asked about how they perceived the hardness of the session in the Börg scale. This dialog situation is handled exactly in the same way as in the Conversation mode. The strategy followed to provide oral feedback is the following:

- The user is informed about the progress of the session three times: when they have completed the first quarter, when they are half way through, and when they have a quarter left.
- The user is cheered up between the progress stages described.
- The user is warned about their pace in case their heart rate is out of the heart rate reserve boundaries. However, there is a 30 second window before the user can be warned about to avoid a constant bombarding of messages insisting them to speed up or slow down.

Every time the current series session is over, the training moves to the next session by increasing the series number. In case the series number had reached its maximum, it is set to 0 and the repetition number is increased. If the number of repetitions had reached its maximum, the training session is over.

4.2 | *Conversational* mode

The second mode corresponds to the interaction with the system when it plays the role of a coach before and after each exercise session. In this mode, the coach follows an architecture which is similar to traditional multimodal dialog systems (see Figure 1), enhanced by an emotion recognition

module. BLE devices are not mandatory in this mode. Thus, the user can talk to the agent through the smartphone or can also wear the heart rate sensor so that the system can check how their heart rate changes when preparing or remembering the exercise sessions.

In a setting before the exercise, the coach and the user can discuss the schedule for the next set of exercises, and the coach can provide appropriate support depending on the user state. For instance, if the agent finds that the user is very stressed, it may try to provide some comfort and try to relieve the stress by planning a schedule with smaller challenges so that the user can experience some progress more easily. At any time after completing an exercise, users can switch the system to *Conversational* mode and describe how they felt. This information is then used to update the user model and the training schedule.

In order to handle the dialog between the application and the user, both TTS and ASR modules are used. The basic idea for the dialog is that the application will prompt a message to the user (generally, asking something) and the user will reply to that message. This cycle will be repeated for a number of times until one of the possible actions is completed (create, change, or cancel a workout). The dialog strategy is as follows:

- The *playPrompt* method attempts to utter a message identified by its utterance ID via the TTS object. A *dialogTag* variable tracks the current stage of the dialog.
- When the message has been generated, the *processTtsResults* method is called. This method calls *listenToAnswer* to listen to the user's response using the ASR object.
- When the recognition is done, the *processAsrResults* method is called. This method retrieves the best answer from the list of possible responses *nextDialog*.
- *nextDialog* is the main method of this activity. The *dialogTag* variable allows to know the current state of the dialog and detect if something went wrong (e.g., the recognizer could not understand anything, a specific piece of information was not found, multiple choices were made instead of one, etc.).

The conversation module is implemented in the Conversation activity. This module allows users to set up a new workout if they do not have a currently active workout or to either change or cancel their workout if they have a currently active workout. These processes are implemented as a spoken dialog between the application and the user:

- When the user decides to change a workout or one of its series, the screen will show a summary of the workout or the series, which will be updated while the user interacts with the application.
- When the user is setting up a new workout, the screen will show the user's diseases (or all the available diseases if they suffer none) and the available exercises for the workout when the system asks the user to chose the value of those parameters.

The operation of this mode is much more complex than that of all the previously explained modules, since the interaction is based only on a dialog. The application guides the user through the process by requesting them to select among a set of choices at every step. As the user answers the application's questions, the process moves on and the user can achieve different objectives such as creating, changing, or canceling their workout.

First of all, when the Conversation activity is launched, the user's prescribed workouts are retrieved from the database repositories. At this point, there is a fork in the conversation flow:

- If the user has an active workout, they can change it or cancel it.
- If the user does not have an active workout, they have to previously set up one.

The selected disease(s) is used to determine the adequate workout for the user. Next, the possible exercises of the selected workout are retrieved and shown on the screen to ask the user to choose an exercise from the list. Finally, the user is asked to choose the days of the week they would like to train, based on the minimum frequency allowed by the workout chosen. The workout prescribed to the user is initialized with the minimum value for all its parameters, and those values will be increased as the user progresses. Users can also change an existing workout. If the user chooses this option, they will be first asked about whether they want to make a general change in the workout (frequency, days of training) or a more specific change in one of the series (intensity, Börg, and duration).

One of the most prominent functionalities of this module is the analysis of the user's emotional states to generate the feedback provided by the application. Once the application has finished recognizing the user's response, it is sent to a sentiment analysis web service. The web service chosen for the practical implementation and assessment of this functionality is MeaningCloud[7], a text analytics and semantic processing software that provides different functionalities related to topic extraction, text classification, sentiment analysis, and text clustering. Given the text to be analyzed, the sentiment analysis provided by this software includes the following information:

[7] https://www.meaningcloud.com/developer/sentiment-analysis/doc. [Accessed December 2018].

- *status*: success of the extraction process
- *model*: model used in the evaluation, along with the language
- *score_tag*: polarity found (strong positive, positive, neutral, negative, strong negative, without sentiment).
- *agreement*: whether there is agreement between the sentiments detected in the text (agreement, disagreement).
- *subjectivity*: whether the text is objective or subjective.
- *confidence*: confidence of the sentiment analysis of the text.
- *irony*: whether the text contains irony marks (ironic, non-ironic).
- *sentence_list*: list of sentences of the text, individually analyzed.
- *sentimented_entity_list*: list of entities of the text, characterized by a certain polarity.
- *sentimented_concept_list*: list of concepts of the text, characterized by a certain polarity.

The information retrieved by the *score_tag* and the agreement are used by the application to reply to the user depending on the feedback they transmitted. This is particularly important when the user provides a negative feedback (e.g., to encourage users to calm them down and prevent them from quitting).

The series' parameters (intensity, duration) of the user's prescribed workout are updated based on the user's progress through the session, the extent to which they completed the workout, and the feedback provided. The workout will only be updated if the user was able to fully complete the session with a good performance (their heart rate was within the HRR range at least 80% of the time) and its feedback was positive. It is recommended to increase the duration first and, then, the intensity. The frequency is updated depending on the updates of the series' parameters. In this regard, if the frequency can be incremented by 3 days/week at most, then the prescribed workout's frequency will be updated every time the user has been able to increase the intensity and the duration of the series by a third of the maximum value of that parameter.

The results module shows a final overview of the session and allow users to provide feedback about the session verbally (see Figure 4. When the user finishes the training session, they are redirected to this activity, which shows a set of bar charts. Each bar chart represents the evolution of the user's heart rate throughout each of the series of the session. In order to convey more complete information, the color of the bars in the chart depends on the user's heart rate.

In addition, the application interacts orally with the user to provide a summary about the session. This summary includes information regarding the extent to which the user completed the series of the session. Then, the user is asked to provide some feedback about the session, which is used to determine the user's overall feeling about the session. Finally, the user's prescribed workout parameters are updated depending on the user's progress.

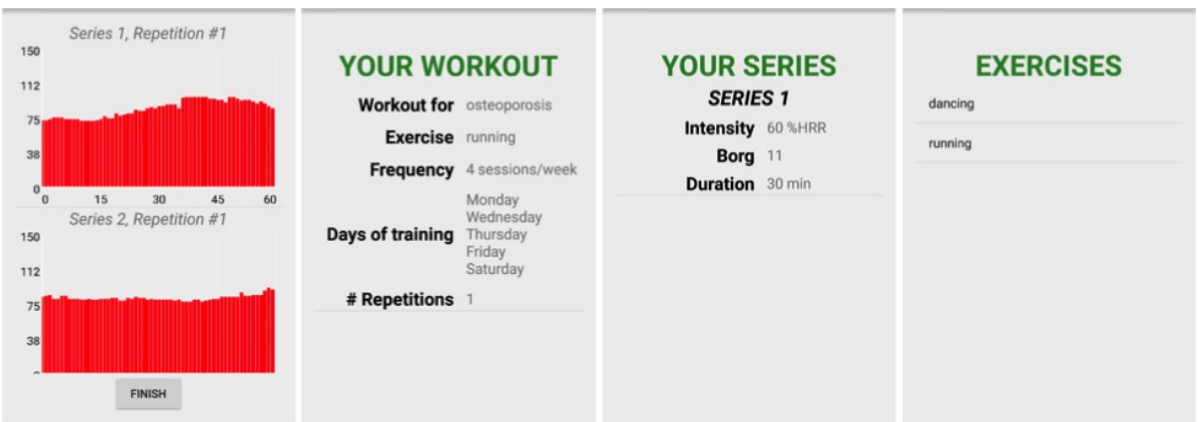

FIGURE 4 Results and conversation screens showing a summary and options of the user's workout or series

5 | PRELIMINARY EVALUATION

It is very difficult to find procedures and measures unanimously accepted by the scientific community for the evaluation of multimodal systems (Hofmann et al. 2014; Metze, Wechsung, Schaffer, Seebode, & Möller 2009). Task-oriented multimodal conversational systems are usually evaluated in terms of interaction parameters and subjective judgments (Callejas & López-Cózar 2008; Möller et al. 2006). Interaction parameters include the technical robustness and core functionality of the system components as well as system performance measures such as task completion rate; whereas subjective usability evaluations estimate features like naturalness and quality of the interactions, as well as user satisfaction reported in questionnaires and interviews. We have completed an evaluation of the developed multimodal conversational coach that is focused on two main aspects. The first one is usability, measuring the benefits of multimodal interaction compared to unimodal interfaces (only spoken or tactile interaction). The second main aspect is the assessment of the technical quality and active ageing potential with professionals and caregivers.

5.1 | Usability assessment of the different input and output modalities

The methodology used to evaluate the multimodal interaction with the conversational coach is based on the work presented in (Metze et al. 2009; Wechsung, Engelbrecht, Kühnel, Möller, & Weiss 2012), which points out that the usability assessment of a multimodal system requires the evaluation of the different interaction modalities. To do this, we have developed the assessment questionnaire shown in Table 3, which is based in standard questionnaires like AttrakDiff (Hassenzahl, Burmester, & Koller 2003) and SASSI (Hone 2014).

TABLE 3 Questionnaire designed for the usability assessment of the developed conversational coach

Previous experience using multimodal interfaces	
Q1. Assess on a scale of 1-5 your previous experience using voice interfaces.	(1 = "Low", 5 = "High")
Q2. Assess on a scale of 1-5 previous experience using multimodal interfaces.	(1 = "Low", 5 = "High")
Understanding of user responses	
Q3. How well did the system understand your responses?	Never, Seldom, Sometimes, Usually, and Always
Q4. How well did the system understand your feelings when given your feedback after the sessions?	Never, Seldom, Sometimes, Usually, and Always
Understanding of system responses	
Q5. How well did you understand the system messages?	Never, Seldom, Sometimes, Usually, and Always
Interaction rate	
Q6. In your opinion, the interaction rate was...	Very slow, Slow, Suitable, Fast, Very fast
Difficulty level using the system	
Q7. Indicate the difficulty level of the system.	Very difficult, Difficult, Normal, Easy, Very easy
Presence of errors	
Q8. Have you noticed errors during the interaction?	Never, Seldom, Sometimes, Usually, and Always
Certainty of what to do at each moment	
Q9. Was it easy to decide what to do after each system turn?	Never, Seldom, Sometimes, Usually, and Always
Active ageing objective	
Q10. Are the suggested workouts well-adapted to your condition?	Not at all adapted; Somewhat not adapted; Just fine; Somewhat adapted; Very well adapted
Q11. Are the training sessions helping you getting fit and actively age?	Nothing at all, A little, Sometimes, Usually, and Very Much
Global satisfaction with the system	
Q12. In general, are you satisfied with the system?	Very dissatisfied; Dissatisfied; Satisfied; Quite satisfied; Very satisfied

As it can be observed, the questionnaire consists of 12 questions. Each question has 5 possible answers from which only one is selected. The main aspects that are evaluated are users previous experience using multimodal interfaces, the degree to which the user finds that the coach understood them and they understood the system, the perceived interaction rate, the perceived difficulty level of the interaction with the system, the presence of errors, the certainty of the user of what to do at each time, the objective of promoting healthy and active ageing, and the global level of satisfaction with the system. The evaluation was completed by 7 native Spanish users (3 men, 4 women, aged 93 to 64, avg. age 68.6).

Firstly, the users evaluated the coach application using the tactile mode. Subsequently, they evaluated the application using the spoken interaction. Finally, they evaluated the multimodal mode combining the use of traditional interfaces (screen and keyboard) with spoken interaction. The users freely chose the actions to be taken and modules and sub-modules to be accessed during the evaluation.

Table 4 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.

TABLE 4 Results of the usability assessment of the multimodal coach (For the mean value M: 1=worst, 5=best evaluation)

Question	Avg. value	Max. value	Min. value	Std. deviation
Q1	3.34	4	1	1.28
Q2	3.10	4	1	1.36
Q3 (tactile mode)	4.76	4	5	0.35
Q3 (oral mode)	4.66	4	5	0.38
Q3 (multimodal mode)	4.85	4	5	0.36
Q4 (tactile mode)	4.81	4	5	0.21
Q4 (oral mode)	4.83	4	5	0.43
Q4 (multimodal mode)	4.94	4	5	0.12
Q5 (tactile mode)	4.62	4	5	0.28
Q5 (oral mode)	4.10	3	5	0.69
Q5 (multimodal mode)	4.78	4	5	0.31
Q6 (tactile mode)	3.64	3	5	0.78
Q6 (oral mode)	3.49	2	5	1.03
Q6 (multimodal mode)	4.61	3	5	0.45
Q7 (tactile mode)	4.13	3	5	0.54
Q7 (oral mode)	3.86	3	5	0.67
Q7 (multimodal mode)	4.73	3	5	0.39
Q8 (tactile mode)	4.60	4	5	0.41
Q8 (oral mode)	3.89	2	5	0.96
Q8 (multimodal mode)	4.74	4	5	0.47
Q9 (tactile mode)	4.02	3	5	0.47
Q9 (oral mode)	4.43	3	5	0.62
Q9 (multimodal mode)	4.81	4	5	0.35
Q10 (tactile mode)	4.42	4	5	0.22
Q10 (oral mode)	4.18	3	5	0.44
Q10 (multimodal mode)	4.46	4	5	0.49
Q11 (tactile mode)	4.64	4	5	0.21
Q11 (oral mode)	4.56	4	5	0.41
Q11 (multimodal mode)	4.77	4	5	0.41
Q12 (tactile mode)	4.44	4	5	0.50
Q12 (oral mode)	3.64	3	5	0.74
Q12 (multimodal mode)	4.88	4	5	0.32

With respect to the extent to which the users feel that the system understood them, it is possible to see that the recognizer had a very good performance. Tactile mode was perceived as more accurate than oral as expected, but the oral mode was punctuated very high by the users, with 4.66 over 5 respectively. As can be observed, users felt that the system understood them better with the multimodal than with the tactile mode (4.85 over 5). In the three interaction modalities, users agree that the application was able to extract the feelings behind theirs feedbacks, given that the sentiment analysis software chosen for the practical implementation of the conversational coach is very accurate and this has greatly contributed to these results (4.94 over 5 in the multimodal mode).

The users also felt that the system responses were comprehensible, specially with the tactile and multimodal modes. The results of the oral mode were lower probably due to the quality of the synthesized voice. We used the standard Spanish voice provided by Android. The results were higher when using better synthesizers like the voices of IVONA TTS. With regard to this point, the multimodal mode was punctuated with 4.78 points over 5.

Regarding the interaction rate, it was found adequate in most cases, though in some cases the participants reported they were expecting barge-in mechanisms in the oral mode. Occasionally, the interaction rate was considered slow if a big transaction data was done in the database repositories (e.g., to record a whole session), but overall, the interaction rate was considered rather fluid. The multimodal mode was found very useful to for the users to interact with the system at a pace which is adequate for their needs, as they could switch between interaction modes. This is in consonance with the fact that the multimodal mode has reported the maximum perceived easiness of use (4.61 over 5).

Generally, the users have not perceived errors during their interactions in the tactile mode, while in the oral and multimodal modes more errors were detected though they did not imply a fail in the interaction and in every case they could complete the task. In particular, the participants reported that the multimodal mode was more useful than the oral mode for reporting errors to the users (4.74 over 5).

About the certainty of what to do at each moment of the interaction, participants felt more security in the multimodal mode, which has also received the better punctuation in the overall satisfaction as it brings more flexibility to the user (4.81 over 5).

Regarding the assessment of the active ageing objective, users overall considered that the proposed workouts were oriented to improve their physical condition and adapted to their age ranges and diseases. Some of the users also proposed extending the application to include additional activities, and the possibility of configuring warm up, break and cool down workouts to help the user start exercising slowly and then to return to calm progressively. The multimodal mode has reported the maximum perceived achievement of this objective (4.46 and 4.77 over 5 in questions Q10 and Q11 respectively).

Also, from the interactions of the users with the coach we completed an objective evaluation considering the following interaction parameters:

- Question success rate (SR). This is the percentage of successfully completed questions: system asks - user answers - system provides appropriate feedback about the answer;
- Confirmation rate (CR). It was computed as the ratio between the number of explicit confirmations turns and the total of turns;
- Error correction rate (ECR). The percentage of corrected errors.

The results of this evaluation (Table 5) shows that the developed application could interact correctly with the users in most cases, achieving a question success rate of 96.34%. Additionally, the approaches for error correction by means of confirming or re-asking for data were successful in 93.11% of the times when the speech recognizer did not provide the correct answer.

TABLE 5 Results of the objective assessment of the interaction parameters

SR	CR	ECR
96.34%	12.80%	93.11%

5.2 | Assessment with professionals and caregivers

An additional evaluation of the application has also been completed with the participation of 6 professionals and caregivers, who rated the naturalness, technical quality and potential of the system for active ageing. The questionnaire shown in Table 6 was defined for the evaluation. The

responses to the questionnaire were measured on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The experts were also asked to rate the system from 0 (minimum) to 10 (maximum) and there was an additional open question to write comments or remarks.

TABLE 6 Questionnaire designed for the evaluation of the multimodal coach with professionals and caregivers

Technical quality
TQ01. The coach offers enough interactivity
TQ02. The coach 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 coach is helpful
TQ07. The coach is attractive
TQ08. The coach reacts in a consistent way
TQ09. The coach provides adequate verbal feedback
TQ10. The coach provides adequate non-verbal feedback
TQ11. The feedback provided by the system improves understanding
TQ12. The system encourages continuing using it after errors
Active ageing potential
AA01. The system fulfills the objective of promoting active ageing
AA02. The proposed activities are relevant for this objective
AA03. The design of the activities was adequate for different kinds of users
AA04. The activities support significant improvements
AA05. The coach promotes a long-term relationship with the users
AA06. The coach complements the workouts without distracting or interfering with them

The results of this questionnaire are summarized in Table 7. As can be observed from the responses to the questionnaire, the satisfaction with technical aspects was high, as well as the perceived potential to promote active ageing and healthy habits. The application was considered attractive and adequate and the experts felt that the system is appropriate and the proposed activities relevant. The global rate for the system was 8.8 (in the scale from 0 to 10).

Although the results were very positive, in the open question the caregivers also pointed out desirable improvements. One of them was to make the system listen constantly in spoken exercises instead of using the push-to-talk interface. However, we believe that this would cause many recognition problems, taking into account the range of exercises. Also, although they considered the contents useful and attractive and its feedback adequate. They also proposed creating a website that would allow users to propose new exercises and new workouts. Similarly, they could suggest to include workouts for diseases not yet considered in the application. Newly-proposed workouts and exercises should be reviewed by a training expert before being incorporated to the application in order to ensure that they are safe to be completed by elderly people.

6 | CONCLUSIONS AND FUTURE WORK

As life expectancy increases, there is a need to find ways to support healthy ageing. This is the purpose of active ageing initiatives that aim to foster healthy behaviors in the population. In this paper we have presented a proposal to build a multimodal coaching conversational agent based on m-health and sentient computing. This agent is developed in Android and supports two modes. On the one hand, an *exercise* mode that processes information from sensors dynamically producing run-time advice based on the user state, the characteristics of the user, the knowledge about the exercise, the history of the interaction and the user's achievements. On the other hand, the agent features a *conversational* mode that offers coaching dialogs between the user and the system, before and after the exercise sessions, with emotion recognition from voice and optionally from the biosignal captured by sensors.

TABLE 7 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
AA01	5/5	5.00	0.00
AA02	4/5	4.87	0.16
AA03	4/5	4.33	0.41
AA04	4/5	4.31	0.43
AA05	4/5	4.62	0.37
AA06	4/5	4.74	0.26

The sensors are widely available in general sport shops at a cheap price, and do not pose special requirements to be used with the coach. Also, the *conversational* mode can be used without sensors, so that the coaching agent can be used with just a smartphone. Given the flexibility of the architecture proposed to develop the multimodal conversational coach, different kinds of modules and devices (e.g., ASR, TTS, sensors, sentiment analysis toolkits, etc.) can be easily integrated and replaced. In addition, as the system considers both continuous and intervallic sessions for the workout's design, it can also be easily adapted to promote other types of activities related to active ageing.

For future work we plan to carry out an evaluation of the multimodal coach with a larger sample of end users to assess their satisfaction with the system, but also to obtain a richer knowledge base for the training sessions with feedback from expert personal trainers. We want to also extend the use of machine learning and Big Data techniques to extract valuable knowledge from the results generated by the use of the application. This could be extremely useful to refine the workouts, based on how the elderly people actually complete them, if they are able to finish them, their perceptions, etc.

References

Allen, J., Ferguson, G., Blaylock, N., Byron, D., Chambers, N., Dzikovska, M., … Swift, M. (2006). Chester: towards a personal medication advisor. *Journal of Biomedical Informatics*, *39*(5), 500–513.

Andre, E., Bevacqua, E., Heylen, D., Niewiadomski, R., Pelachaud, C., Peters, C., … Rehm, M. (2011). Emotion Oriented Systems. The Humaine Handbook. Cognitive Technologies. In (pp. 585–608). Springer Verlag.

Bhatia, M., & Sood, S. (2017). A comprehensive health assessment framework to facilitate IoT-assisted smart workouts: A predictive healthcare perspective. *Computers in Industry*, *92-93*, 50–66.

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

Bickmore, T., Puskar, K., Schlenk, E., Pfeifer, L., & Sereika, S. (2010). Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers*, *22*, 276–288.

Bickmore, T. W., Schulman, D., & Sidner, C. (2013). Automated interventions for multiple health behaviors using conversational agents. *Patient Education and Counseling*, *92*(2), 142–148.

Bickmore, T. W., Schulman, D., & Sidner, C. L. (2011). A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *Journal of Biomedical Informatics*, *44*(2), 183–197.

Bourret, R., & Bousquet, J. (2013). An integrated approach to telemonitoring noncommunicable diseases: best practice from the European

innovation partnership on active and healthy ageing. *World Hospitals and Health Services*, *49*(3), 25–28.

Callejas, Z., & López-Cózar, R. (2008). Relations between de-facto criteria in the evaluation of a spoken dialogue system. *Speech Communication*, *50*(8-9), 646–665.

Cavazza, M., de la Cámara, R. S., & Turunen, M. (2010). How Was Your Day? a Companion ECA. In *Proc. of AAMAS'10* (pp. 1629–1630). Toronto, Canada.

Chiu, C., & Liu, C. (2017). Understanding older adult's technology adoption and withdrawal for elderly care and education: mixed method analysis from national survey. *European Journal of General Practice*, *19*(11), 374.

Christophorou, C., Kleanthous, S., Georgiadis, D., Cereghetti, D., Andreou, P., Wings, C., … Samaras, G. (2016). ICT services for active ageing and independent living: identification and assessment. *Healthcare Technology Letters*, *3*(3), 159–164.

Cvecka, J., Tirpakova, V., Sedliak, M., Kern, H., Mayr, W., & Hamar, D. (2015). Physical activity in elderly. *European Journal of Translational Myology*, *25*(4), 249–252.

Dalgaard, L., Gronvall, E., & Verdezoto, N. (2013). Mediframe: A tablet application to plan, inform, remind and sustain older adults' medication intake. In *Proc. of ICHI'2013* (pp. 36–45). Philadelphia, PA, USA.

Dimitrova, R. (2013). Growth in the intersection of e-health and active and healthy ageing. *Technology and Health Care*, *21*(2), 169–172.

Farzanfar, R., Frishkopf, S., Migneault, J., & Friedman, R. (2005). Telephone-linked care for physical activity: A qualitative evaluation of the use patterns of an information technology program for patients. *Journal of Biomedical Informatics*, *38*, 220–228.

Ferreira, S., Sayago, S., & Blat, J. (2017). Older people's production and appropriation of digital videos: an ethnographic study. *Behaviour and Information Technology*, *36*(6), 557–574.

Fleg, J. (2012). Aerobic exercise in the elderly: a key to successful aging. *Discover Medicine*, *13*(70), 223-228.

Foster, C., Richards, J., Thorogood, M., & Hillsdon, M. (2013). Remote and web 2. 0 interventions for promoting physical activity. *Cochrane Database of Systematic Reviews*, *2*, 1–84.

Gates, N., Karim, S., Rutjes, A., Ware, J., Chong, L. Y., March, E., & Vernooij, R. (2016). Computerised cognition-based interventions for maintaining cognitive function in cognitively healthy people in midlife. *Cochrane Database of Systematic Reviews*, *7*, 1–17.

Gellish, R., Goslin, B., Olson, R., McDonald, A., Russi, G., & Moudgil, V. (2007). Longitudinal modeling of the relationship between age and maximal heart rate. *Medicine and Science in Sports and Exercise*, *39*(5), 822-829.

Giorgino, T., Azzini, I., Rognoni, C., Quaglini, S., Stefanelli, M., Gretter, R., & Falavigna, D. (2004). Automated spoken dialogue system for hypertensive patient home management. *Journal of Medical Informatics*, *74*, 159–167.

Golbidi, S., & Laher, I. (2012). Exercise and the cardiovascular system. *Cardiology Research and Practise*, *2012*, 1–15.

Griol, D., Callejas, Z., López-Cózar, R., & Riccardi, G. (2014). A domain-independent statistical methodology for dialog management in spoken dialog systems. *Computer Speech and Language*, *28*(3), 743–768.

Gronvall, E., & Verdezoto, N. (2013). Beyond self-monitoring: Understanding non-functional aspects of home-based healthcare technology. In *Proc. of UbiComp'13* (pp. 587–596). Zurich, Switzerland.

Harper, R., Nicholl, P., McTear, M. F., Wallace, J. G., Black, L.-A., & Kearney, P. (2008). Automated Phone Capture of Diabetes Patients Readings with Consultant Monitoring via the Web. In *Proc. of ECBS'08* (p. 219-226). Sydney, Australia.

Hassenzahl, M., Burmester, M., & Koller, F. (2003). Mensch & computer 2003. interaktion in bewegung. In (pp. 187–196). Vieweg+Teubner Verlag.

Henriquez-Camacho, C., Losa, J., Miranda, J., & Cheyne, N. (2014). Addressing healthy aging populations in developing countries: unlocking the opportunity of eHealth and mHealth. *Emerging Themes Epidemiology*, *11*(1), 136.

Hofmann, H., Hermanutz, M., Tobisch, V., Ehrlich, U., Berton, A., & Minker, W. (2014). Evaluation of In-Car SDS Notification Concepts for Incoming Proactive Events. In *Proc. of IWSDS'14* (pp. 102–112). Napa, USA.

Hone, K. (2014). *Usability measurement for speech systems : SASSI revisited* (Tech. Rep.). London, UK: Brunel University.

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

Jongstra, S., Beishuizen, C., Andrieu, S., Barbera, M., van Dorp, M., van de Groep, B., … Richard, E. (2017). Development and validation of an interactive internet platform for older people: the healthy ageing through internet counselling in the elderly study. *Telemedicine journal and e-health*, *23*(2), 96–104.

Karki, A., Savel, J., Sallinen, M., & Kuusinen, J. (2013). Ethicted (evaluation process model to improve personalised ICT services for independent living and active ageing)-future scenario. In *Proc. of ICHI'13* (pp. 159–162). Vilamoura, Portugal.

Kaufman, H. (2012). From where we sit: augmented reality for an active ageing european society. *Journal of Cybertherapy and Rehabilitation*, *5*(21), 35–37.

Keijser, W., Manuel-Keenoy, E. D., D'Angelantonio, M., Stafylas, P., Hobson, P., Apuzzo, G., … Senn, A. (2016). DG connect funded projects on information and communication technologies (ICT) for old age people: beyond Silos, CareWell and SmartCare. *Journal of Nutrition, Health and*

Aging, *20*(10), 1024–1033.

Knowles, L., Skeath, P., Jia, M., Najafi, B., Thayer, J., & Sternberg, E. (2016). New and future directions in integrative medicine research methods with a focus on aging populations: a review. *Gerontology*, *62*(4), 467–476.

Laranjo, L., Dunn, A., Tong, H., Kocaballi, A., Chen, J., Bashir, R., … Coiera, E. (2018). Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, *25*(9), 1248–1258.

Lattanzio, F., Abbatecola, A., Bevilacqua, R., Chiatti, C., Corsonello, A., Rossi, L., … Bernabei, R. (2014). Advanced technology care innovation for older people in italy: necessity and opportunity to promote health and wellbeing. *Journal of the American Medical Directors Association*, *15*(7), 457–466.

Lee, E., Han, S., & Jo, S. (2017). Consumer choice of on-demand mhealth app services: context and contents values using structural equation modeling. *International Journal of Medical Informatics*, *97*, 229–238.

Lee, G., Kim, H. K., Jeong, M., & Kim, J. (2015). *Natural Language Dialog Systems and Intelligent Assistants*. Springer.

Leite, I., Pereira, A., Castellano, G., Mascarenhas, S., Martinho, C., & Paiva, A. (2012). Modelling empathy in social robotic companions. *Advances in User Modeling*, *7138*, 135–147.

Lionis, C., & Midlov, P. (2017). Prevention in the elderly: A necessary priority for general practitioners. *European Journal of General Practice*, *23*(1), 203–208.

Martínez-Tomás, R., Fernández-Caballero, A., & Ferrández, J. M. (2014). Intelligent monitoring for people assistance and safety. *Expert Systems*, *31*(4), 343–344.

McTear, M. F., Callejas, Z., & Griol, D. (2016). *The conversational interface: Talking to smart devices*. Springer.

Metze, F., Wechsung, I., Schaffer, S., Seebode, J., & Möller, S. (2009). Human-computer interaction. novel interaction methods and techniques. In (pp. 75–83). Springer-Verlag.

Möller, S., Englert, R., Engelbrecht, K., Hafner, V., Jameson, A., Oulasvirta, A., … Reithinger, N. (2006). MeMo: towards automatic usability evaluation of spoken dialogue services by user error simulations. In *Proc. of Interspeech'06* (pp. 1786–1789). Pittsburgh, USA.

Mooney, K., Beck, S., Dudley, W., Farzanfar, R., & Friedman, R. (2004). A computer-based telecommunication system to improve symptom care for women with breast cancer. *Annals of Behavioral Medicine Annual Meeting Supplement*, *27*, 152–161.

Muellmann, S., Forberger, S., Mollers, T., Zeeb, H., & Pischke, C. (2018). Effectiveness of ehealth interventions for the promotion of physical activity in older adults: a systematic review. *Systematic Reviews*, *108*, 93–110.

Payr, S. (2010). Closing and closure in human-companion interactions: Analyzing video data from a field study. In *Proc. of ro-man'10* (pp. 476–481). Viareggio, Italy.

Pfeifer, L., & Bickmore, T. (2010). Designing Embodied Conversational Agents to Conduct Longitudinal Health Interviews. In *Proc. of IVA'10* (pp. 4698–4703). Philadelphia, USA.

Pieraccini, R. (2012). *The Voice in the Machine: Building computers that understand speech*. MIT Press.

Pinto, E. (2007). Blood pressure and ageing. *International Journal of Medical Informatics*, *83*(976), 109-114.

Poole, E. (2013). HCI and mobile health interventions: How human-computer interaction can contribute to successful mobile health interventions. *Translational Behavioral Medicine*, *3*(4), 402–405.

Preschl, B., Wagner, B., Forstmeier, S., & Maercker, A. (2011). E-health interventions for depression, anxiety disorder, dementia, and other disorders in older adults: a review. *Journal of Cyber Therapy and Rehabilitation*, *4*(3), 371–385.

Rehrl, T., Geiger, J., Golcar, M., Gentsch, S., Knobloch, J., Rigoll, G., … Wallhoff, F. (2013). The Robot ALIAS as a Database for Health Monitoring for Elderly People. In *Proc. of AAL'13* (pp. 414–423). Berlin, Germany.

Reijnders, J., Geusgens, C., Ponds, R., & van Boxtel, M. (2017). Keep your brain fit! Effectiveness of a psychoeducational intervention on cognitive functioning in healthy adults: a randomised controlled trial. *Neuropsychological Rehabilitation*, *27*(4), 455–471.

Richard, E., Jongstra, S., Soininen, H., Brayne, C., van Charante, E. M., Meiller, Y., … Andrieu, S. (2016). Healthy ageing through internet counselling in the elderly: The hatice randomised controlled trial for the prevention of cardiovascular disease and cognitive impairment. *BMJ Journal*, *6*, E010806.

Robbins, T. D., Keung, S. N. L. C., & Arvanitis, T. N. (2018). E-health for active ageing. a systematic review. *Maturitas*, *114*, 34–40.

Santos, V., Reis, A., & Barroso, J. (2013). A game to improve memory in elderly. In *Proc. of TISHW 2016* (p. A19). Vila Real, Portugal.

Sixsmith, A., Meuller, S., Lull, F., Klein, M., Bierhoff, I., Delaney, S., & Savage, R. (2009). SOPRANO - An Ambient Assisted Living System for Supporting Older People at Home. In *Proc. of ICOST'09* (pp. 233–236). Tours, France.

Smith, C. D. (2005). *A comparison of the effects of anaerobic and aerobic exercise on mood: A dissertation*. Spalding University.

Tiedemann, A., Rissel, C., Howard, K., Tong, A., Merom, D., Smith, S., … Sherrington, C. (2016). Health coaching and pedometers to enhance physical activity and prevent falls in community-dwelling people aged 60 years and over: study protocol for the Coaching for Healthy AGEing (CHAnGE) cluster randomised controlled trial. *BMJ Open*, *6*(5), 1–8.

van Dyk, S., Denninger, S. L., & Richter, A. (2014). The many meanings of active ageing. confronting public discourse with older people's stories. *Recherches sociologiques et anthropologiques*, *44*(1), 97–115.

Wechsung, I., Engelbrecht, K., Kühnel, C., Möller, S., & Weiss, B. (2012). Measuring the Quality of Service and Quality of Experience of multimodal human-machine interaction. *Journal on Multimodal User Interfaces*, *6*(1-2), 73–85.

WHO. (2002). *World Health Organization. Active ageing: a policy framework*. WHO/NMH/NPH/02.8. Available from: https://www.who.int/ageing/publications/active_ageing/en/. [Accessed December 2018].

WHO. (2018). *World Health Organization. The Global Network for Age-friendly Cities and Communities*. Available from: https://www.who.int/ageing/gnafcc-report-2018.pdf?ua=1. [Accessed December 2018].

Young, S. (2011). Cognitive user interfaces. *IEEE Signal Processing Magazine*, *27*(3), 128–140.

AUTHOR BIOGRAPHY

David Griol. Dr. David Griol obtained his Ph.D. degree in Computer Science from the Technical University of Valencia (Spain) in 2007. He has also a B.S. in Telecommunication Science from this University. He is currently visiting lecturer at the Department of Computer Science in the Carlos III University of Madrid (Spain). He has participated in several European and Spanish projects related to natural language processing and dialog systems. His research activities are mostly related to the development of statistical methodologies for the design of spoken dialog systems. His research interests include dialog management and optimization, corpus-based methodologies, user modeling and simulation, adaptation and evaluation of spoken dialog systems and machine learning approaches.

Jose Manuel Molina. Dr. Jose Manuel Molina is Full Professor at Universidad Carlos III de Madrid, Spain. He joined the Computer Science Department of the same university in 1993. Currently, he coordinates the Applied Artificial Intelligence Group (GIAA). His current research focuses on the application of soft computing techniques (NN, evolutionary computation, fuzzy logic, and multiagent systems) to radar data processing, air traffic management, and e-commerce. He (co)authored up to 75 journal papers and 200 conference papers. He received a degree in telecommunications engineering from the Technical University of Madrid in 1993 and the Ph.D. degree from the same university in 1997.

Araceli Sanchis. Dra. Araceli Sanchis is Full Professor of Computer Science at Universidad Carlos III de Madrid (UC3M) since 1999. She received her Ph.D. in Computer Science from UPM in 1990 and in physical Chemistry from Complutense University of Madrid in 1994. She has a B.S. in Chemistry (1991) from the Complutense University of Madrid. She has been vicedean of the Computer Science degree at UC3M and, currently, she is head of the AI CAOS group (Grupo de Control, Aprendizaje y Optimización de Sistemas), based on machine learning and optimization. She has published over 110 journal and conference papers mainly in the field of machine learning applications.