Multimodal Technologies for Seniors: Challenges and Opportunities[1]

Cosmin Munteanu, Albert Ali Salah

October 2016

---

[1] This is the uncorrected author proof. Please cite this work as: Munteanu, C., A.A. Salah, “Multimodal technologies for seniors: Challenges and Opportunities,” S. Oviatt, B. Schuller, P. Cohen, D. Sonntag, G. Potamianos, A. Krüger (eds), The Handbook of Multimodal-Multisensor Interfaces, Morgan Claypool, 2017.

0.1 Introduction

This chapter discusses interactive technologies in the service of seniors. Adults over sixty-five form one of the largest and most rapidly growing user groups in the industrialized society. Interactive technologies have been steadily improving in their ability to provide practical support for these users and their daily challenges. These applications range from supporting low-literacy adults with mobile touch and spoken language interfaces [Munteanu et al. 2014], to interactive robotic fitness coaches for seniors [Görer et al. 2016]. The goal of this chapter is to present an in-depth analysis of the needs of seniors, as well as to investigate how multimodal and multisensor technologies are used to address these needs.

In Section 0.2 of this chapter, we structure the needs of older adults according to a hierarchy of human needs that emphasizes self-actualization [Maslow 1954]. Maslow's theoretical prism helps us to identify a broad spectrum of potential opportunities for researchers and practitioners who design multimodal-multisensor applications. We present a number of example applications in Section 0.3. These examples are illustrative of how interactive technologies, especially ones that leverage multiple modalities (e.g. speech, gestures, touch) and sensor data (e.g. location and accelerometers available on consumer phones) can provide a basis for meeting the needs of older adults. Section 0.4 briefly surveys future directions of multimodal-multisensor research that has potential for developing valuable new *assistive living technologies*[1]. Section 0.5 is a discussion of both design and implementation challenges, followed with conclusions in Section 0.6, and a list of supplementary Focus questions to aid in further understanding of essential content. We hope that this broad overview will encourage both new and established researchers to explore how current or future multimodal and multisensor interfaces can better support one of the largest but most under-represent demographic groups.

0.2 Senior Users and Challenges

One challenge facing both gerontology and *gerontechnology* research is the difficulty in defining the target population. The age to be considered "elderly" varies from 50 [Ziefle et al. 2008] to 88 [Lepicard and Vigouroux 2010]. The characteristics of this population vary widely, but involve physical abilities, cognitive capabilities, education, digital literacy, etc. These characteristics distinguish the elderly from a general population, and may require specialized HCI solutions.

In this section we review the challenges faced by older adults with respect to their specific needs. We organize these according to Maslow's hierarchy of human needs [Maslow 1954] depicted in Figure 1. In gerontechnology, this taxonomy has been used as a framework for considering the potential utility of healthcare technology for older adults [Thielke et al. 2011]. We deal with each level in separate subsections, starting from the more basic needs.

---

[1] See Table 1 for a glossary of terms used in this chapter.

Table 1 Glossary of key terms.

Alzheimer's	disease is "the most common type of dementia that causes problems with memory, thinking and behavior" [Alzheimers' Association 2016]. Alzheimer's usually progresses over time and may lead to a patient to become incapacitated in performing routine activities.
Ambient technology	is technology integrated into the environment, to endow it with sensing and acting capabilities.
Assistive living technology	encompasses technology to facilitate daily life of older adults, to improve their independence for longer periods.
Exergames	are computer-based (video) games in which the player's body acts as a game controller and that are designed to engage the player in physical activity or exercise.
Gerontechnology	is a field of research and development concerned with designing technology and environments to support older adults' independent living, social participation, good health, comfort, and safety.
Living lab	is a home-like laboratory environment, equipped with many sensors and monitoring systems, where subjects actually live during the course of an experiment.
Participatory design	allows the users of a system to provide direct input during the design of the system.
Self-actualization	is an individual's need to achieve their full potential with respect to their abilities: "even if all needs are satisfied, we may still often (if not always) expect that a new discontent and restlessness will soon develop, unless the individual is doing what they are fitted for." [Maslow 1943].
Smart homes	are living accommodations that allow their occupants to interact with computing technology embedded in the construction of the living space.
Social isolation	is "a state in which the individual lacks a sense of belonging socially, lacks engagement with others, has a minimal number of social contacts and they are deficient in fulfilling and quality relationships" [Nicholson 2012]. Social isolation can be measured quantitatively, which is in contrast to "loneliness" – a state of mind often experienced by older adults, regardless of their social isolation status [Perissinotto et al. 2012].

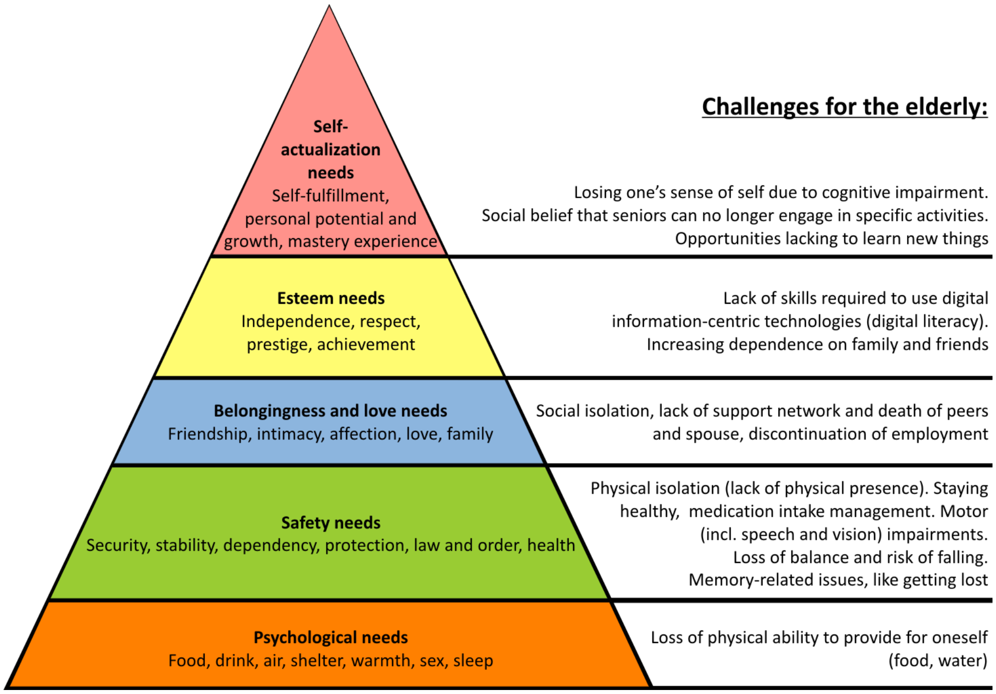


Figure 1 Maslow's hierarchy of basic human needs [Maslow 1954], and the corresponding challenges for the elderly. The most basic needs (i.e. physiological) are placed at the bottom.

0.2.1 Physiological needs

Maslow's hierarchy defines physiological needs, such as food, drink, air, and shelter, as the most basic of all human needs. It is possible to address some of these needs through technological interventions. Assistive living technologies are developed to facilitate the daily life of older adults and to improve their independence for longer periods. We describe this area in more detail in Section 0.3.3.

The first challenge for older adults is meeting fundamental needs, such as sustenance. A study of community-dwelling older adults (e.g. living in retirement or nursing homes) showed that 15% of residents are at risk of malnutrition [Bollwein et al. 2013]. Feeding is one of the core physiological needs that multimodal technologies can address outside of clinical / healthcare environments. We illustrate the use of multimodal technology for addressing this problem with a couple of examples.

Our first example is Brian — an assistive interactive robot that guides community-dwelling older adults in daily tasks like eating [McColl et al. 2013]. Brian has a speech-based interface, but through motions of its head, torso and arm, as well as by modulations of their speed, it

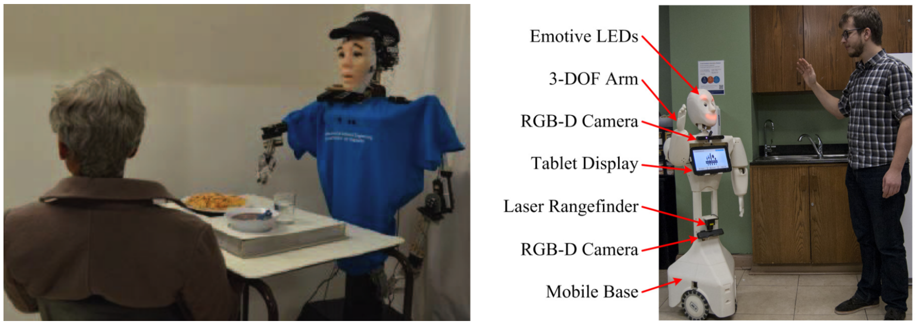


Figure 2 Left: Brian 2.1 socially assistive robot for the elderly and cognitively impaired (Source: [McColl et al. 2013]). Right: The Casper multimodal robotic kitchen helper. (Source: [Bovbel and Nejat 2014], Video at: https://youtu.be/noSJ9qWt_f0)

can convey emotional gestures as well [McColl and Nejat 2014]. This second modality, used simultaneously with speech, increases engagement and compliance with the recommended eating process. The actions of the robot are divided into groups like "orienting" and "encouraging," each of which is performed with a different emotional tone. Brian is also a multisensor system, in that it uses 2D, depth and infrared cameras to perceive the person interacting with it, as well as several weight sensors to perceive the food items on the tray (see Figure 2, left).

A similar example is the meal-assistance robot proposed by [Tanaka et al. 2014], which helps older adults with the mechanics of lifting items from a food tray. Gaze is used to control this robot, so that individuals who struggle with physical impairments can use it.

In addition to interactive robots, prosthetic and orthotic technologies are developed. Prosthetics such as exoskeleton robotic arms replace missing limbs, whereas orthotics enhance and modify existing neuromuscular and skeletal systems. An example orthotic is the commercially available Soft Extra Muscle system [Nilsson et al. 2012], which is a multisensor glove to strengthen the hand (see video at https://www.youtube.com/watch?v=Blg9iK1P8gs for a use case).

0.2.2 Safety needs

Safety needs for older adults include health, daily living, preventing accidents, and mobility. Health is the safety need most addressed by technology, especially through multimodal-multisensor processing [Alemdar and Ersoy 2010]. The range of interactive assistive technologies addressing health includes self-monitoring devices, ambient sensors, and telehealth interfaces.

In recent years, the advent of mobile and wearable technologies, as well as affordable, low-power, yet accurate sensors have created significant new opportunities. There is an increase in commercial mobile health (mHealth) technologies, some specifically marketed to seniors (e.g. Jawbone's UP and BeClose). Advances in machine learning made it possible to extract additional relevant signals, such as the heart rate, from consumer-grade sensors in mobile phones [Han et al. 2015]. However, clinical research is often critical of the accuracy claims of many mass-market mHealth devices and applications [Kumar et al. 2013], and calls into question how the health/safety needs of vulnerable users are met by such technologies.

Health is intrinsically connected to assistive living technologies for older adults [Abowd et al. 2002]. Older adults require most help with activities that are inherently private, such as personal hygiene [Czarnuch et al. 2011], which requires technology design that supports personal dignity and esteem (see Section 0.2.4).

A major assisted living application area is fall detection and prevention, as falls are one of the most important threats to older adults' safety [Kannus et al. 2005]. There are a multitude of clinical, medical, and pharmacological-based approaches to reducing the incidence of and mitigating the effects of falls. However, medical research indicates that long-term implementation of these approaches in assisted living or in-home contexts is quite difficult [Tinetti et al. 2004]. The extensive survey conducted by [Farshchian and Dahl 2015] illustrates that the research focus has been mainly on sensor technology. This has included (1) improving the processing capabilities for more accurately detecting and predicting signs of imminent falls [Belshaw et al. 2011], (2) using fewer and less intrusive ambient sensors, (3) leveraging sensor capabilities of mobile phones [Alemdar and Ersoy 2010, Mellone et al. 2012], and (4) utilizing low-power wearable devices [Bertolotti et al. 2016].

A suitably placed omnidirectional camera can detect falls in a home environment [Demiröz et al. 2014], but with multiple sensors, it becomes possible to predict them. [Bourennane et al. 2013] proposed a system that combined ambient sensors (light, infrared, and magnetic), with a pressure sensor mounted on the user's bed and with a wearable RFID (radio frequency identification) sensor attached to the user's back. A supervised learning algorithm for combining data from these sensors into a behavioral model was proposed to predict falls. Such a system can also anticipate and monitor other events, such as nocturnal restlessness or immobility, changes in amount of movement, and such behavioral deviations. A similar system has been proposed by [Castillo et al. 2014], combining multiple stationary and wearable sensors as afforded by consumer-grade devices: video cameras, accelerometers, and portable location trackers. The machine learning algorithm proposed by the authors relies predominantly on video analysis, which is complemented by accelerometer and location data. The authors show that such a multisensor approach can lead to the detection of up to 80% of falls. Finally, [Yavuz et al. 2010] proposed a multimodal interface for requesting assistance after a fall. The system leverages the multisensor capabilities of a smartphone to detect the fall, creates an emergency alert, and allows its user to manage the alert through the phone's touch interface.

One of the challenges faced by fall prevention systems is noisy and inaccurate data, and multisensor approaches are useful for improving robustness [Patel et al. 2012]. A second important issue is users' acceptance of omnipresent and intrusive technologies that eliminate privacy, as well as constantly reminding them of their limitations [Hawley-Hague et al. 2014]. Multiple modalities and sensors may exacerbate this issue. A third challenge is that of integration of wearable and *ambient technologies*. An example is the HIPPER project, which helps patients to perform exercises after a hip surgery in their home environment [Aicha et al. 2016]. This is a truly multidisciplinary research space, as illustrated by many novel approaches taken to address the challenges of fall prevention strategies, such as gamification and *participatory design* for seniors [Uzor et al. 2012].

In many cases, this type of research is still preliminary and the results are often obtained in artificial settings, using simulated data sets, or with a very small number of participants. The system described by [Bourennane et al. 2013] has been evaluated through a long-term deployment in a hospital room, with a single senior user suffering from *Alzheimer's disease*. The system of [Castillo et al. 2014] was trained with an existing dataset, and tested on actors staging falls. The HIPPER project was conducted in three *living labs*, which provide ecologically valid conditions, but are expensive to set up and maintain.

0.2.3 Love and belonging

Social needs such as family, sense of belonging, and friendship, are difficult to maintain as people age, yet are considered to be of significant importance [British Columbia Ministry of Health 2004]. Unfortunately, several studies indicate high incidences of *social isolation* [Nicholson 2012] or loneliness [Perissinotto et al. 2012] among seniors. However, even a simple telecommunication technology that connects an older adult to loved ones can have a great impact [Dishman 2004].

Online communication and social media technologies have a significant potential to reduce social isolation for older adults [Chen and Schulz 2016], and thus address their love and belonging needs to a certain extent. Technologies like online video chat have been shown to be effective for maintaining a sense of love and belonging in long-distance relationships for young and mid-age adults [Neustaedter and Greenberg 2012]. However, commercially-available communication technologies pose significant barriers for adoption by older adults [Barnard et al. 2013]. Some of these barriers can be overcome by making such technologies multimodal, which allows older adults to interact more easily through the modality with which they feel most comfortable. For example, [Teixeira et al. 2012] developed a communication appliance that supports touch, gestures, voice, and typing as input modalities. A usability study with ten older adults revealed a preference for combining touch and voice when operating the appliance.

Multisensor technologies can facilitate other forms of non-verbal communication. In the SnowGlobe application, designed by Visser et al. [2010], one globe is placed in the house of the older adult, and the other in the house of a family member. The SnowGlobe (Figure 3) uses motion and light sensors to detect the movements of the remote users and glows brighter as the

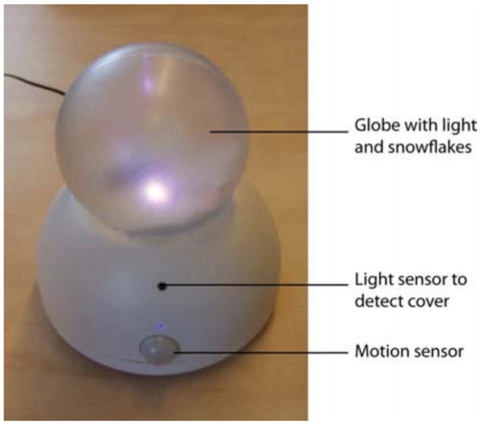


Figure 3 The SnowGlobe multisensor, non-verbal communication appliance. (Source: [Visser et al. 2010]).

remote user moves more actively (in his/her room), thus conveying the feeling of presence. It can also display a subtle call for attention; if the user shakes the globe, the remote globe paired with it will blink, which acts as a more direct notification of the user thinking about the other person. By using subtle background cues to orient others, this social awareness application clearly addresses a need of the older adults not to be a burden to their families. Another such example is the i-Pot tea-kettle sold in Japan by Zojirushi. It incorporates a wireless transponder and sends a message to designated recipients over the Internet every time it is used, as well as regular e-mails about the most recent usage, thus signaling well-being.

0.2.4 Esteem

A person's need for esteem is often defined in relation to current societal, cultural, and economic expectations and norms. One of the main correlates of esteem is the sense of independence [Sato and Cameron 1999]. At the core of ensuring independence is the need to design technologies that support older adults' information-centric needs. For example, [Aly and Munteanu 2016] argue that older adults' lack of digital literacy leaves them reliant on others for understanding and managing complex information (e.g. about their health). Assistive technologies can provide the proper balance between giving older adults control over the information seeking and managing processes and providing the targeted support they need. In this context, the usability of assistive technologies is key to facilitate digital literacy for older adults [Borges and Sinclair 2008], and multimodality can provide the support for this [Munteanu et al. 2011].

As with the other needs discussed earlier in this section, the need for esteem and independence is interconnected with the need for belonging, family, and friendship. As surveyed in [Neves et al. 2015], issues like low digital literacy can affect older adults' connection to their loved ones and to their peers, especially since current social and economic shifts have created families living at great distances. Strengthening social ties and leveraging older adults' existing social support can also increase their independence and self-esteem when coping with information challenges [Carmien and Fischer 2008].

0.2.5 Self-actualization

Interactive technologies that empower the elderly to function independently and help them to assume meaningful responsibilities and active job roles can contribute to their personal growth and *self-actualization*. In a longitudinal study conducted in a group care facility, seniors who were given responsibility for tasks like watering their own plants had better health profiles and better longevity outcomes compared to seniors for whom plants were cared for by staff [Mallers et al. 2014]. One implication of this research is that multimodal-multisensor interface design for seniors should balance active user control over the system using input modalities like speech, touch, typing, and writing, with more passive sensor- and camera-based activity tracking. Passive or fully-automatic technologies can undermine a person's perceived control and motivation. In the long term, balanced multimodal-multisensor interfaces can contribute to seniors' perceived independence, sense of purpose, motivation to remain engaged, as well as to their health.

A study conducted by Conference Board of Canada [2010] showed that the increasing technological and information demands in the workplace may raise additional barriers for older adults struggling with maintaining meaningful employment later in life. Unfortunately, most current assistive technologies only provide direct and restricted assistance with respect to a specific impairment. An example would be multimodal reading aids [Harrison 2004], which often fail to report long-term success, as surveyed by Siegenthaler et al. [2010]. A possible solution is to design self-calibrating systems that can learn certain parameters from the input provided by the users, thereby adapting to the user as the usage behavior changes [Passerini and Sebag 2015].

0.3 Specific Application Areas

In this section we review several application areas within the space of gerontechnology and assistive technologies for older adults.

0.3.1 Socially assistive robotics

Socially assistive robots (SAR) designed for seniors mostly focus on monitoring or else helping the elderly with their daily lives [Heerink 2010, Pineau et al. 2003]. This assistance is primarily through social interaction, which predominantly requires speech- and gesture-based interfaces, as well as direct input for efficiency [Feil-Seifer and Mataric 2005]. SAR involves a socially-oriented human embodiment, and as such, goes beyond simple devices that provide assistance for rehabilitation, mobility and education [Colombo et al. 2007]. When designing robotic assistance for older adults, social embodiments and interaction are very important to maintain long-term engagement [Wada et al. 2005]. This section focuses on applications involving physical exercise, and illustrates some practical aspects of SAR related to multimodal interaction research.

Most existing systems developed for elderly physical exercise do not involve robots at all, but rather screen-based interfaces [Barnes et al. 2009, Sucar et al. 2009]. An embodied conversational agent (ECA), or a similar 3D avatar displayed on a screen, can provide a realistic visualization of the target exercise. However, physical and tangible SAR embodiments have the advantage of being more engaging than screen-based ECA interfaces. SAR applications currently are proliferating largely due to the increasing availability of robotic platforms, and also recent developments with the Robotic Operating System (ROS) [Quigley et al. 2009] that includes a standard message-passing interface to facilitate multimodal integration on robots that use it. Fasola and Mataric [2013] have contrasted user responses to 3D physical and 2D screen-based virtual robots in an exercise scenario with the elderly, and showed that the physical robot was rated as more engaging, enjoyable, and a better exercise partner. Similarly, Lopez Recio et al. [2013] showed that in a physiotherapy scenario, real robots provoked better mimicry responses compared to simulated robots. The embodiment aspect is closely related to motivation, which is a key element in sustaining attention and physical exertion over a long time. In [Fasola and Mataric 2012], a robotic exercise coach was proposed for chair aerobics, and the authors evaluated the motivational aspects of this scenario extensively. One motivational strategy the authors used was providing numeric feedback on task success, which “gamified” the experience and made it more engaging. Secondly, the robot avoided giving negative feedback.

The dominant input modality to a social robot is speech, provided that the application language has sufficient support. For under-resourced languages, spoken SAR interactions are typically kept simple, and operated with restricted vocabularies. With the introduction of cheaper depth cameras, visual input also advanced rapidly, supporting real-time gestural input. This input mode now can be sensed and processed more easily and affordably [Mollaret et al. 2016]. Additionally, social robots often use color cameras to detect faces and expressions. They can incorporate tactile sensors to recognize touch [Yang et al. 2015]. These modalities can be used to analyze users’ affective states, and then provide appropriate non-verbal signals such as backchannel nods and social smiles to improve interaction. Fusion of modalities is not performed at the data level; for instance there are no integrated systems yet that understand speech audio-visually. However, rule-based fusion of modalities like speech and gestures at the dialogue management level is achieved [Stiefelhagen et al. 2007].

The synchronization of feedback and back-channel signals is crucial for interaction with older adults, because of the increased risk of confusion due to deteriorating perceptual channels. Language comprehension is known to decrease with age-related reduction in working memory capacity [Kemper and Mitzner 2001]. Consequently, corresponding non-verbal signals become more important. The increased cognitive load of interacting with novel technologies also increases the risk of confusion. Multimodality allows the seniors to choose the most accessible and usable modality for them. However, if the synchronization of multimodal feedback delivered by the robot is not well co-timed, seniors become confused and communication failures ensue [Görer et al. 2016].

The evaluation of SAR requires the assessment of the system in both physical and social dimensions [Sim and Loo 2015]. For the latter, personality alignment, degree of empathy, and behavioral adaptation can be measured [Tapus and Mataric 2008]. For measuring interaction issues, the number of help requests and errors are evaluated, as well as vocal and facial expressions of confusion and frustration. We provide a detailed design walk-through example in Section 0.5.2.

0.3.2 Social connectedness and participation

As we have detailed in Section 0.2, older adults' basic needs such as health and safety are often interconnected, and solutions require their active social participation. Since challenges like physical and cognitive disabilities can also undermine social connectedness, it can be difficult to avoid a downward spiral. Interactive technologies can contribute to older adults' sense of independence [Neves et al. 2013], and have been successfully used to help older adults overcome health-related barriers [Dadlani et al. 2010].

An example solution that was also designed to address the complex adoption factors discussed in Section 0.5.1 is the InTouch application [Neves et al. 2015], illustrated in Figure 4. This system provides practical support for maintaining a sense of connectedness to seniors' loved ones. InTouch was designed as an asynchronous replacement for real-time video communications and photo/text messaging. The main features of InTouch are: sending/receiving of images, audio, or video, receiving (but not sending) text messages, and sending of one pre-defined "I'm thinking of you" message. It has a non-language interface based on icons, swipe gestures, tapping, and voice message input, which requires no typing. In one of the few longitudinal evaluations with "oldest old" adults (i.e. over 80), the study established that social, attitudinal, physical, digital literacy, and usability factors contributed to the adoption of communication technologies.

One of the ways of connecting people is to engage them in joint activities, such as reading together. For older adults, reading together with young family members can become an important activity. The People in Books multimodal system proposed by Follmer et al. [2012] blends video, audio, and textual interaction to visually immerse remote participants, such as an elderly person with a grandchild, in a book's storyline. Attarwala et al. [2013]

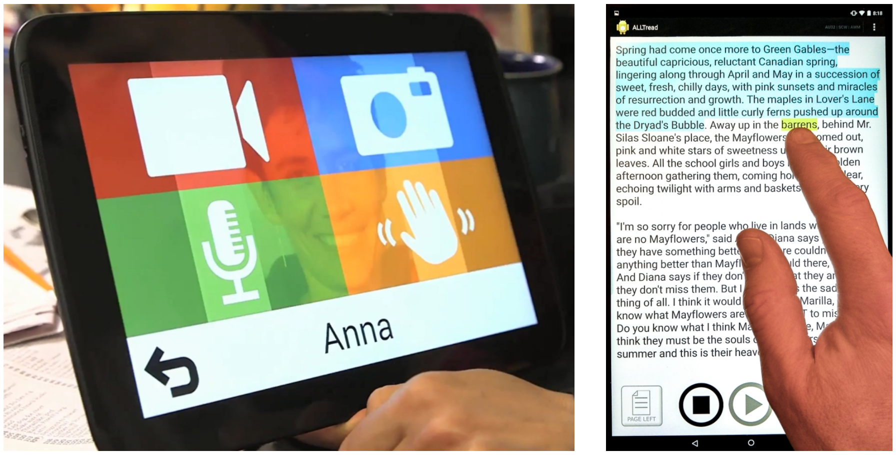


Figure 4 Left: The InTouch [Neves et al. 2015] asynchronous communication application for seniors (Source: provided by authors). Right: The ALLT e-reader ([Attarwala et al. 2013] and http://cosmin.taglab.ca/share/ALLT/video) supporting asynchronous collocated "reading together" (Image source: provided by authors).

introduced the ALLT application, which helps older adults read together with their younger family members in the same place at different times. This system supports younger family members' efforts to create audio e-books to be read by their older relatives, as illustrated in Figure 4. Accurately-synchronized audio/text enables the older adult to later recreate the experience of reading together with a family member by playing back the recording, while following the text. The synchronization is achieved by combining audio processing with tracking the finger, dynamically adapted to a user's natural reading speed.A controlled study with 20 young participants showed that the multimodal combination of finger tracking and intelligent adaptation resulted in users' recording at their optimal reading speed and thus unburdening them from having to constantly follow the text.

Multimodal technologies to address social isolation can also be deployed in an assisted living setting. For example, Teixeira et al. [2012] proposed the Living Home Centre application (Figure 5). They have implemented several multimodal extensions to a *smart home* environment, such as speech, touch, and 3D gestures. These allow older adults to interact with a custom-built messaging, calendaring, and social media application. These extensions were evaluated through a typical usability study [Rogers 2011] in which ten older adults were asked to perform tasks such as sending messages or setting reminders. Speech and touch were the most efficient and preferred modalities when accessing the social communication services provided by the Living Home Center. These findings are applicable not only to normally-ageing seniors, but also to those affected by age-related impairments or dementia, for which

Figure 5 The Living Home Centre Application [Teixeira et al. 2012]. (Image source: [Pires et al. 2012].)

the use of multimodal interaction has been proven effective at increasing social communication [Prange et al. 2015, Waycott et al. 2014].

0.3.3 Assisted living technologies

Ambient assisted living (AAL) is a vision of older adults living in smarter homes, and subsequently retaining their independence for longer periods. The investment in the smart home[2] technology is assumed to pay off considering the financial burden of maintaining older adults in a specialized care facility, especially as longevity increases in the population. Early AAL applications focused on increasing the safety and security of the seniors (e.g. by turning cookers off automatically), providing care functions (e.g. automatic medication notification), and convenience-related functions (e.g. automatic light switches) [van Berlo 2002]. Newer applications in this area include automatic fall detection, automatic detection and monitoring of cognitive and physical illnesses, cognitive assitive tools (orthotics) such as software-based personal reminder systems, and applications for social connectedness [Rashidi and Mihailidis 2013].

The possibility of equipping smart homes with many sensors has been a common solution for activity recognition, to detect presence, falling, and wandering off premises [Alemdar and Ersoy 2010]. One advantage of multimodal systems is their ability to enable "mode-switching", which supports users in changing to another modality if the current one is misrecognized by the system [Perry et al. 2004]. For a fuller description of the major error handling advantages of multimodal systems, see [Oviatt 2002].

In an extensive review of the domain, Solaimani et al. [2013] remark that the main design principles for smart homes are flexibility, reliability, scalability, and non-intrusiveness, all

---

[2] There is another, more commercial use of the term *smart home*, which just implies that the electronic appliances in the home are controlled by a single panel. We use *smart home* to depict a home environment that actively monitors its inhabitants and adapts to them in the process.

of which are beneficial for seniors with impairments. Flexibility results from implementing different interfaces for the same task, and requires extra design effort and higher cost [Kieffer et al. 2009]. However, this effort is often justified. Since seniors' status changes more rapidly than that of younger adults, system functionality and features need to be adapted to support them optimally.

Demiris and Hensel [2008] investigated a number of smart home technologies for assisting older adults. Their survey shows that a lot of research goes into functional monitoring applications (e.g. collection and analysis of data pertaining to activity levels, motion, gait, meal intake, and activities of daily living), and for safety monitoring (e.g. detection of fire and gas leaks, automatic light switches, fall detection), followed by monitoring of physiological signals (e.g. pulse, respiration, body temperature, blood pressure, bladder and bowel output) and providing cognitive support (e.g. medication reminder, verbal assistance for appliances). They remark that the focus is shifting from a data mining perspective to one that targets empowering the elderly and caregivers. Advances in wearable computing and wireless communications support this development, simultaneously promising potential solutions to the lack of interoperable and affordable systems for individuals, which seems to be the major problem of the field [Memon et al. 2014].

Technological healthcare interventions serve the ambient assisted living vision of "aging in place," helping the elderly with basic needs, but also with healthy living and with informal caregiving [Salah et al. 2015]. Privacy is the primary concern in ethics of assistive technology designed for older adults, followed by issues of autonomy, obtrusiveness, acceptability, affordability and safety [Zwijsen et al. 2011]. Older adults often object to technology if they feel that it will diminish the actual social contact they will receive [Görer et al. 2016, Rashidi and Mihailidis 2013]. If the technology is perceived to be impersonal, inappropriately timed, socially insufficient, or embarrassing, it may not be accepted [Dishman 2004]. An example is reported by Görer et al. [2016], where several elderly users rejected the assistance of a NAO robot on the grounds of it being "childish," and not suitable for their age.

0.3.4 Access to information

Most information today is available only in digital, online form. Older adults report feeling overwhelmed and disenfranchised when accessing online information [Aly and Munteanu 2016]. Declining cognitive and visual abilities play a role in this. However, as demonstrated by Teixeira et al. [2014, 2012], multimodal interfaces have the potential to remove some of the usability barriers faced by adults when searching for information online.

Access to information involves applications that help older people to navigate and search the Internet, as well as facilitating access to various media. For example, the GUIDE platform [Coelho and Duarte 2011] is an interface that allows older adults to access information on smart TVs in a multimodal way. This includes speech, gestures, face recognition, touch on tablet surface, and input keys on TV remote. The interface adapts to users' actions or context;

Figure 6 The SimSensei multimodal virtual kiosk [DeVault et al. 2014]. (Image source: [Morency et al. 2015]. Video at: https://youtu.be/ejczMs6b1Q4.)

e.g. it automatically lowers the TV volume when users select speech as input, or increases font size when users remove their glasses. It is also capable of storing a user profile of multimodal input and output preferences. The GUIDE platform was evaluated through a usability study with 17 older participants who were asked to perform TV setup and operation tasks. This revealed that users prefer to interact with the system using multiple modalities at the same time, and that users' characteristics, abilities, and interaction preferences varied widely [Coelho et al. 2011].

Accessing and interacting with information can be facilitated by smart virtual agents. An example is the SimSensei Kiosk [DeVault et al. 2014, Morency et al. 2015], initially designed to detect anxiety, depression, or stress disorders (Figure 6). It exhibits multimodal, human-like behavior to complement the speech-based interaction with users. A usability evaluation with 350 participants of various ages found that the addition of behavior cues to the virtual agent enhanced the perception of naturalness and human-like interaction. This made participants more willing to share health information with the agent. Analysis of users' facial and gestural cues as additional input modalities created more natural and comfortable dialogues [Gratch et al. 2013] .

0.3.5 Personal assistants

Human caregivers are costly. Personal assistant (PA) systems provide services to older adults, to enable them to stay socially more engaged and connected [Hämäläinen et al. 2015], or to be deployed within the context of a home environment [Pires et al. 2012]. Most commercially-available PAs are limited in their multimodal capabilities, typically offering a rudimentary combination of touch and (predominantly) speech input [Cohen et al. 2016]. Nevertheless, PAs have now became widely available commercially, especially in mobile form (e.g. Siri [Apple Inc. 2016], Amazon's Echo [Amazon Inc. 2016]), with recent reports of increased market interest in their adoption by older adults [O'Brien 2016]. These systems work best when fine-tuned specifically for elderly users. For instance the AALFred Personal Life Assistant [Hämäläinen et al. 2015], which offers touch, gesture, and speech interaction, is trained with speech collected from senior users. This considerably improves its use by the elderly compared to standard speech recognisers tuned for younger adult speech.

Assisting older adults with their medication management is one of the most common applications of personal assistants, and several authors have proposed or designed such systems to be multimodal. For example, the S4S Medication Assistant [Ferreira et al. 2013] combines speech input with a depth measurement to infer how far the user is and to adjust display size (Figure 7). The system uses all available multimodal input channels to also learn about a user's context and history in order to provide them with the relevant health advice. The system was designed through an approach combining personas and scenarios, and it was evaluated through usability inspection sessions with subject matter experts. The evaluation was complemented by a task-based think-aloud usability session that asked four users to complete several scenarios expected of a medication assistant, followed by qualitative interviews focusing on users' perception of the multimodal interaction. The evaluation of the system focused on the usefulness of medication information and on reminder features. However, the authors reported that users perceived the combination of speech input with context-aware visual information favorably, although they indicated the need for more accurate speech processing [Ferreira et al. 2014].

The literature shows speech to be a particularly useful single interaction modality for older adults, due to its ease of use, but also due to its efficiency for content creation. However, as we will discuss in Section 0.5.1, multimodality is essential in facilitating the adoption of personal assistants, especially for older adults affected by cognitive decline [Yaghoubzadeh et al. 2013]. This has also been demonstrated by Looije et al. [2010], who conducted a study with 24 older adults in the context of inducing behavior change for management of diabetes. The study employed an anthropomorphized cat-shaped assistant that supports speech input/output, as well as proximity and touch sensors that allow it to mimic human-like social behavior. Users were asked to perform daily activities specific to managing their condition. A within-subject evaluation was conducted, comparing the multimodal robot with a virtual speech-based agent (also depicted as a cat) and with a desktop-based text-only interface performing the same functions. A qualitative analysis of video recordings of participant interactions and qualitative questionnaires revealed that the multimodal interaction lead to higher engagement and elevated conversational behavior of the participants, while the speech input lead to shorter, more effective interactions. The multimodal interface also caused the participants to be more trusting of the behavior-altering advice provided by the assistant.

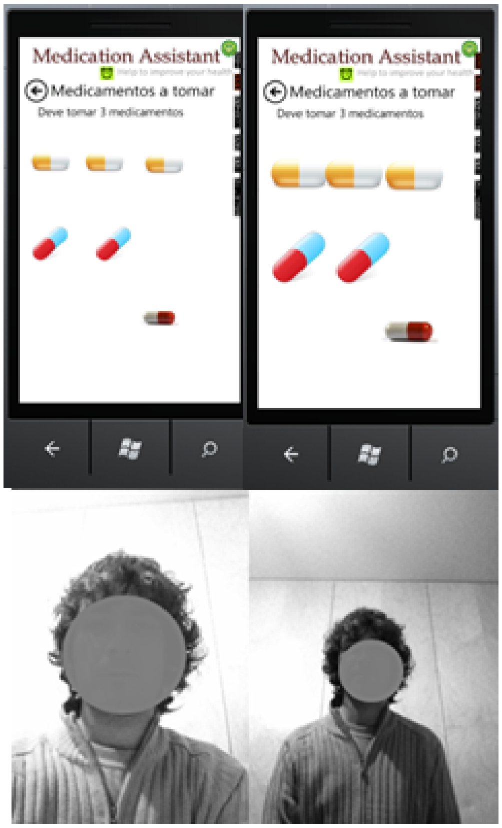


Figure 7 The S4S Medication Assistant [Ferreira et al. 2014].

0.4 Available Multimodal-Multisensor Technologies

Several research projects and studies bring convincing evidence that multimodal-multisensor technologies have reached the maturity level that makes them beneficial to older adults. In this section we review representative examples, as well as commercial applications.

In the field of social robotics, prominent examples include Aldebaran's Pepper [Aldebaran Inc. 2016], which is available commercially, Willow Garage's PR2 [Willow Garage Inc. 2016] or Cynthia Breazeal's Jibo [Jibo Inc. 2016]. For these robots, simple gaze-orientation behavior and pre-coded backchannel signals like head nods and blinks are used for improving the quality of interaction, although automatic speech perception is not sufficiently developed yet to "have a functional conversation". Other assistive robots are dedicated to simply providing emotional comfort to their users, such as the commercially-available multimodal pet robot Paro [Paro Robots U.S. Inc. 2016], which responds to simple voice commands and to physical touches on its fur. Similar to a pet, Paro emits sounds, turns its hand, and modifies its facial

Figure 8 The Toyota Human Support Robot (HSR) system, planned to coexist with family members in a home environment. (Image source: [Kalogianni 2015], and video at: https://www.youtube.com/watch?v=QoS-40Xe75Q.)

expressions. Several videos available from the manufacturer at http://www.parorobots.com/video.asp illustrate how Paro is employed in seniors' homes in countries such as Japan.

Some larger robotic assistants have reached mature development levels as well. The GiraffPlus project [GiraffPlus 2016] developed an assistive environment with a telepresence robot as one of its main components (video at: https://youtu.be/9pTPrA9nH6E). Similarly, Toyota developed a Human Support Robot (HSR) prototype (see Figure 8), which is controlled by a tablet, and provides audio-visual feedback to the elderly. Its design principles are lightweight and maneuverable design, safe interaction, and simple interface. Integration of robotic assistants and tablets was also proposed in the Robot-Era project [Bevilacqua et al. 2015], which combined three service robots for various tasks (see Figure 9).

Commercialization of smart homes is less frequent than that of assistive robots, as a fully integrated smart home is expensive and difficult to maintain [Chan et al. 2008]. Rashidi and Mihailidis [2013] list 17 smart home initiatives across the world, but these initiatives are all research prototypes. Multi-national initiatives like the EU Ambient Assisted Living Joint Programme increased the coordination in this area by funding many international collaborative projects, and helped integration of different technologies [Busquin 2013].

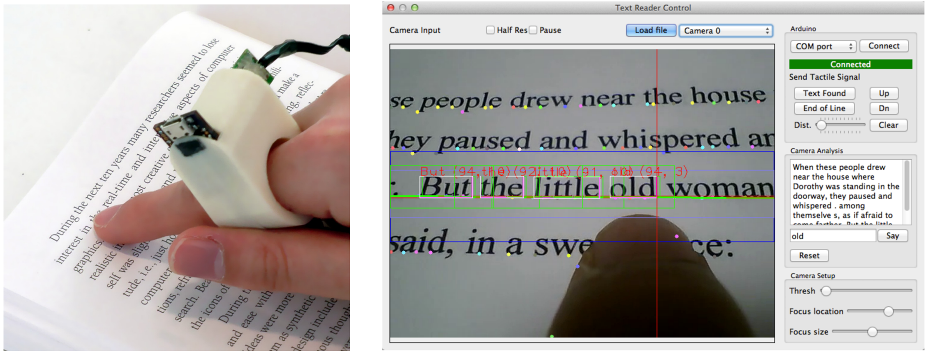
The above examples of commercially-ready or -available assistive technologies often incorporate multimodal interfaces and sensors serendipitously. In many cases, modules handling the individual modalities are not developed from scratch, but adapted from existing tools to the application at hand. Oviatt [2003] has demonstrated that combining modalities in a user interface can significantly reduce the errors in performing the task at hand – a so-

Figure 9 The Robot-Era project [Bevilacqua et al. 2015] combines three robots, to provide assistance to the elderly both indoors and outdoors. (Image source: provided by project authors to [ANSA 2016], and see video at: https://www.youtube.com/watch?v=lv43z8YVQkY.

lution that since became the de facto approach to developing multimodal applications. For commercial solutions, we see functional modules like depth-camera based human tracking or speech processing tools being used across systems with great success, although more strategic approaches to developing fusion-based architectures could further enhance their reliability.

Within gerontechnology, previous research has shown that older adults prefer to communicate with assistive technologies through direct and natural modalities, such as speech for input or visual presence for output [Aylett et al. 2014, Bickmore et al. 2005]. As an input modality to assistive technologies, speech is often combined with touch – e.g. to support activities like map navigation, home media control, and eyes-free interaction while driving. One example is the steering wheel-mounted tablet that allows drivers to interact with a car's environment through voice commands and touch controls [Pfleging et al. 2011]. Several investigations have revealed that older adults would benefit from increased automated assistance while driving [Ziefle et al. 2008]. For an extensive review of recent research in automotive multimodal interaction, see Chapter [Schnelle-Walka and Radomski 2017X].

With respect to accessibility and assistive technologies, speech can be successfully combined with text entry. An example application is an e-reader device that allows low-vision users to read printed text using a small, finger-mounted optical recognition device [Shilkrot et al. 2015] (Figure 10). This system uses RGB-based image analysis for optical tracking,



When these people drew near the house where Dorothy was standing in the doorway, they paused and whispered . among themselves, as if afraid to

old

Figure 10 MIT's FingerReader multimodal wearable device facilitating assisted reading for blind users. (Source: [Shilkrot et al. 2015]. Video at: https://youtu.be/25uPK2POLwc.)

with a heuristic algorithm for detecting finger activity and occlusions. The located text image is processed by an optical character recognition module for tracking the read text. The multimodal output combines audio and haptic feedback: the text is read with a text-to-speech module, and haptic guidance is provided. This device was evaluated in a usability study with three blind users, which reported that directly mapping information to a modality that the users could process was well received.

Such multimodal interfaces are natural candidates to support older adults' needs. For example, Casper – a touch- and speech-enabled robotic kitchen helper [Bovbel and Nejat 2014], is a full-size robot that "holds" a touch tablet at chest level and provides guidance to seniors with meal preparation (Figure 2, right). The tablet supports item selection through touch, while voice is used in lieu of textual entry for functions such as search. The robot also relies on sensors such as RGB-D cameras for tracking users' body postures and distance to the robot. Casper uses a predictive algorithm to locate seniors in their home environments, based on a model of previously observed behaviors and a history of locations. The robot displays basic facial emotions, based on the touch- and speech-based interaction with users.

Progress in RGB-D cameras fueled development of systems with real-time full-body interactions. This has supported the growth of several novel areas such as *serious games* that support workplace skills training, or *exergames* that improve users' health and fitness. Some of these applications are implemented as multimodal interfaces. For example, [Rector et al. 2013] have developed and evaluated a Kinect-based exergame interface that combines speech feedback with body tracking to provide support for blind and low-vision adults in practising yoga. Most social assistive robotics applications (see Section 0.3.1) use full-body tracking in conjunction with speech.

The use of speech as an output modality is somewhat less extensively studied compared to output modalities such as haptics. Significant effort in multimodal research is dedicated to the *input* combination of speech and haptics [Turk 2014]. Of particular interest when designing elder-centred system is the integration of haptics under various conditions (see Chapter [MacLean et al. 2017X]). However, there are several instances of research where audio in form of voice or sonification and haptics are combined as *output*. These are surveyed extensively in Chapter [Freeman et al. 2017X]. An example assistive multimodal system in this category is the BlindNavi mobile application [Chen et al. 2015]. This application uses GPS sensors to accurately locate its users on a map and to provide contextual information to the user such as landmark information, wayfinding details, and available transportation. The multimodal feedback consists of vibration patterns to indicate navigation instructions, and voice to simultaneously deliver the contextual information. A usability field study with four blind users was employed to fine-tune the design of the interface and the amount of information delivered to users, such as the granularity of the landmark details.

A common non-technological assistive device that is always associated with old age is the walking cane. A research prototype built by Megalingam et al. [2015] proposes a "smart" cane that uses an ultrasonic sensor to detect obstacles and to measure their distance. The smart cane provides simultaneous voice and haptic feedback. The obstacle information is reported to the user through voice alerts. The ultrasonic distance measurements are used to infer whether the obstacle is in motion, which is conveyed to the user as vibration patterns in the cane, and the vibration intensity varies with the speed of the obstacle.

0.5 Multimodal Interaction for Older Adults – Usability, Design, and Adoption Challenges

Usability is critical for the adoption of assistive technologies [Venkatesh et al. 2003], and multimodality can increase usability for older adults [Teixeira et al. 2012]. Still, older adults continue to struggle with using multimodal interfaces, especially when rigorous user-centred design processes are not followed with respect to how modalities are combined. We analyze here how multimodality and particularly efforts to more seamlessly combine or integrate modalities may affect the usability of gerontechnology. We then discuss how modalities that are dominant or natural in everyday life are perceived by older adults when used to interact with technology, and how individual differences within this demographic group may affect this perception.

Humans are extremely capable of integrating different modalities during human-to-human interaction. Research such as [Oviatt et al. 2003] showed that seniors have the same integration patterns as younger adults, and that they are able to switch between modalities with similar lags as other age groups. This is called intermodal sequential integration. However, physiological differences of older adults do affect how seniors use multimodal interfaces [Czaja and Lee 2002]. Such differences range from cognitive difficulties [Salthouse 2016, Thomas et al. 2010], to language processing [Kemper and Mitzner 2001], and to visual acuity [Jacko et al. 2002]. Zajicek and Morrissey [2003] found that older users, when using a multimodal online browsing aid, have difficulties processing complex information through speech. They also often prefer to use a single modality that suits better their abilities. Such difficulties may be even more pronounced with advanced age. A study by Neves et al. [2015] of a multimodal tablet-based application showed that the oldest of older adults (80+) found some modalities or combination of modalities too complex to use. These included typical mobile gestures such as swiping and tapping: swiping motions were difficult for users with hand tremors, while the lack of tactile feedback and physical affordances resulted in users hitting the tablet very hard while tapping virtual buttons.

In one of the most relevant studies dedicated to understanding multimodal integration patterns in older adults, Xiao et al. [2003] found that adults aged 66 to 86 that used a speech- and pen-based map interface exhibited substantially more individual differences in multimodal integration patterns compared to younger adults. This indicates that multimodal interfaces designed for seniors could benefit more from *adaptive integration* thresholds. On the other hand, Siek et al. [2005] investigated performance differences between younger and older adults when using a touch- and pen-based personal device, and concluded that with appropriate training time, differences between groups become statistically negligible. Naumann et al. [2010] reported an investigation on multimodal interaction with common mobile interfaces through speech, gestures, and touch controls, and found no difference in performance between younger and older adults. However, the study also revealed that this was due to older adults' being able to select their modality of choice and to use it almost exclusively. The *ability to choose modalities* is obviously an important consideration.

Historically, modalities such as real-life speech show large variations in error rates, and subsequently are challenging for interface design [Huang et al. 2014]. Older adults have been shown to have higher error rates compared to young adults [Vipperla et al. 2008]. Recent advances in deep neural networks significantly improved automated speech recognition [Hinton et al. 2012], as well as gesture recognition [Wu et al. 2016], in terms of speed and accuracy. Such advances diminish the gap between error rates of older and younger adults, and eliminate some of the hurdles facing multimodal integration.

The findings surveyed in this section highlight the need to gain a deeper understanding of the challenges faced by older adults when interacting with a multimodal system, and that the main issues are not engineering, but design and usability problems.

0.5.1 Design Considerations and Approaches

Jian et al. [2014], Neves et al. [2015] propose several recommendations for the design of specific multimodal interface elements in elder-centred multimodal systems. We summarize these in Table 2. Additionally, McGee-Lennon et al. [2011] suggest broader-context design recommendations for increasing the adoption of multimodal assistive technologies that support older adults, extending the recommendations proposed by Naumann et al. [2010] and Teixeira et al. [2012]. We illustrate these in Table 3.

Interactive technologies for older adults have disproportionately focused on how technology can accommodate the physiological or cognitive abilities typically associated with this cohort, while often neglecting non-technological factors [Moffatt 2013]. Neves et al. [2013] found that attitudes may play a role in seniors' adoption of interactive technologies. In a follow-up study, Neves et al. [2015] found that in addition to the physiological and cognitive factors identified in previous research, socio-demographic and cultural factors affect how older adults use multimodal technologies. Examples include incorrect interpretation of icons, perception of some gestures as awkward, and lack of willingness to learn how to operate new technologies.

Table 2 Design guidelines for UI elements of multimodal technologies supporting older adults [Jian et al. 2014, Neves et al. 2015].

User challenge	UI element	Design recommendation
Visual perception	Layout	Simple, clear
Hearing loss	Speech output	Low-pitch, vigorous
Dexterity	Touch controls	Large, regular shape
Dexterity	Weight	Lighter hand-held controls
Attention	Text and icons	Consistent font, colors, icon sizes
Memory	Item selection	Limit of three choices, associated with images and keywords
Strong bias	Icons and gestures	Adaptation to users' cultural preferences
Digital literacy	Multimodal UI affordances	Clear and explicit usage instructions

No single design approach represents a guaranteed solution to balancing the usability issues highlighted here. However, it is clear that early focus on "getting the design right" is important for addressing older adults' diverse needs. Several approaches exists that facilitate this, under the larger framework of User-Centred Design [Preece et al. 2015]. We highlight two approaches particularly suitable for gerontechnology: Participatory Design and Contextual Design, respectively.

Participatory Design [Schuler and Namioka 1993] (PD) involves users at all stages of the design, and elicits their direct input for specifying the design and the functional requirements of a system. PD can provide a rich context for the overall design and development of multimodal technologies, especially for older adults. This includes understanding the needs of older adults [Vines et al. 2012], selecting the appropriate modalities to interact with the

Table 3 Guidelines for the contextual design of multimodal assistive technologies, focusing on increasing their adoption by older adults [McGee-Lennon et al. 2011, Naumann et al. 2010, Teixeira et al. 2012].

User challenge	Interaction context	Design recommendation
Diverse abilities	Need to select the most suitable modality	Personalization by users or their caregivers
Diverse abilities	Need to leverage user's strongest or most preferred modality	Dynamically-adapted multimodal interfaces
Specific habits	Even simple tasks (e.g. reminders) need to fit well with user's personal routines	Adapt to user's temporal and location context
Independence	Support user's need for self-reliance, privacy, and independence	User-initiated interaction with multimodal inputs
Privacy	Older adults more frequently share space with others	Personalization of output modalities to better suit the context of use
Reliability	Users need to be able to rely on their assistive technologies for critical support	Employ well-developed components and rely on complementary modalities to reduce error rates and to increase usability

assistive technology [Nicol et al. 2016], and customizing the features offered by a multimodal assisted living interface [Muñoz et al. 2015].

Contextual Design [Beyer and Holtzblatt 1997] (CD) is an approach based on understanding users and their needs as they perform tasks in their natural environment. This is called *contextual inquiry*. CD is well suited for developing technology for the elderly, as it facilitates the establishment of functional requirements without making explicit the social, economical, personal factors that affect technology use, and their complex interrelations. CD can be applied iteratively, starting from gathering requirements to validating prototypes and to understanding the adoption of the technology. It has been successfully employed for validating technology that addresses needs at multiple levels, as we discussed earlier, such as applications targeting social isolation [Neves et al. 2015, Waycott et al. 2014].

User-Centred Design incorporates other approaches for including the users in the early stages of design. However, as Franz et al. [2015] indicate, not all may be as suitable for the elderly as PD or CD. For example, the *think-aloud* approach, which requires users to vocalize what they are doing and thinking during system usage, may be too challenging, as it requires older adults to perform simultaneous tasks. Psychometric quantitative data collection methods such as the Likert-scale questionnaires may be confusing, as the differences between items are not always evident. The presence of researcher-as-participant may lead the older adults to tailor their responses to impress or please the researcher, thereby affecting the generalizability of the results.

0.5.2 Case Study: Designing and Implementing a Multimodal Assistive Robot

As a case study, we describe here the design of a robotic fitness coach [Görer et al. 2016], schematically depicted in Figure 11. The robot has a full-body sensing module that uses an RGB-D camera, a text-to-speech module to provide speech-based feedback in Turkish, and a motion-based feedback module that uses its human-like joints to illustrate fitness exercises.

Tested for use in an elderly care facility in Turkey, this robot tracks senior users visually and walks them through a set of fitness exercises (see video: https://youtu.be/lbLo3-oIi8o). Since Turkish speech-to-text tools are not sufficiently developed, the robot can not follow verbal commands. However, it provides audio feedback in Turkish, which is complemented by visual demonstrations. The audio is useful for the cases where the robot's gesture palette is not expressive enough. For instance, it lacks hand joints and cannot illustrate the *clenching the hands* exercise visually. Gestures are essential for the cases where the participant has hearing issues. One third of the elderly participants of this study (12 subjects from an elderly care facility, mean age 82.2) were observed to have hearing difficulties. The particular feedback during exercises depends on robot's full-body tracking of the participant, as the robot judges the performance thereby. The robot can perform a restricted set of stretching and relaxation exercises (see [Görer et al. 2013] for more details).

During the design of this system, the requirements elicitation stage established which exercises should be included, but the hearing problems became evident only after the first set of observational studies. A longitudinal user study clarified how the elderly participants perceived the robot, identified cultural issues and allowed fine-tuning of the social interaction. A nursing home was visited, and a professional exercise tutor was consulted to select exercises based on the type, stance pose and appropriateness for mimicking by the robot. The researchers attended regular exercise sessions, and had debriefing sessions with the instructor at the end of these sessions. Corrective and positive feedback are used by the professional tutor, and the robot was similarly designed to provide these two types of feedback. The timing of feedback is carefully engineered. During observation sessions, verbal explanations for each motion were recorded, and their order in the program was noted. Balance and endurance exercises are not selected for robotic tutoring due to risk of falling and heart problems.

Two preliminary studies were performed to test the robot's motion learning and transfer capabilities and the use of the system. The situations that require corrective feedback (such as speed adjustments, amplitude adjustments, mirroring detection, incorrect imitations), as well as empirical parameters (such as thresholds for giving feedback) are determined during these studies, and options are generated for the feedback. A preliminary study was conducted with young participants, followed by a second preliminary study with seniors. At the end of each study, participants filled out a questionnaire in which they indicated their perception of social aspects of the system, and the overall system performance also was assessed. This questionnaire was based on the Game Experience Questionnaire (GEQ) that measures different emotional responses to a game-like experience [Norman 2013]. The authors measured positive and negative affect, flow, immersion and challenge on a 5-point Likert scale.

Interaction-related issues were improved by a final preliminary study in an adult day care center. Video recording and photo caption were not allowed in the nursing home due to privacy concerns. One researcher attended the sessions as an observer and took notes about any unexpected behaviors of the subjects, as well as their facial expressions. Finally, longitudinal tests with seniors were performed to evaluate the system. If video recording were permitted, gaze and attention analysis could have been used for additional evaluation.

During sessions, a verbal explanation is provided by the robot for each exercise, followed by a gesture demonstration, which the elderly participant is supposed to imitate. The robot tracks the participant to assess the imitation of the motions. This is a purely quantitative measure of performance. At the end of the final motion, the robot informs the participant that the session is completed, and gives an overall performance score for the session. The robot has a set of response templates to give corrective or positive feedback. Since automatic speech perception and generation were difficult for the robot, these response templates were supplemented by a Wizard of Oz setup for verbal communication, in which the experimenter could modify or supplement the responses by typing a specific feedback response on the fly. The robot would use the text-to-speech module to vocalize these responses.

The evaluations showed that the system scored high for acceptance as an exercise tutor, but not as a social companion or friend. The primary reason for the latter was determined to be the limited dialogue capabilities of the robot. Some of the subjects stated that they did not want to assume the responsibility of taking care of the robot, especially because it involved technology with a risk of malfunctioning. Some subjects expressed a fear of becoming dependent on the robot. On the other hand, the participants enjoyed the sessions, and were motivated to perform well as indicated by the scores provided at the end of the session.

We note here that there are numerous questionnaires and constructs that can be selected to assess aspects of usability for assistive robots, and the choice of GEQ in the case study puts more emphasis on aspects of flow and affect. A very simple usability assessment tool is the System Usability Scale (SUS), developed by Brooke et al. [1996], which produces a single value between 0 and 100. The Unified Theory of Acceptance and Use of Technology (UTAUT) is one of the most perused theoretical frameworks to assess user acceptance [Venkatesh et al. 2003], and was adapted to social robotics by Heerink et al. [2010] in their Almere model. This model incorporates constructs for perceived ease of use and adapt-

Summary of Design Process: The design begins with **establishing requirements**. The designers work with a professional fitness coach to establish a taxonomy of elderly exercises and to discuss usability [Görer et al. 2013]. This is followed by the **development of submodules** required for two-way interaction, including full-body sensing, text-to-speech, and motion-based feedback. **In situ observational studies** (videos and surveys) are performed in an elderly care facility to evaluate functionality and engagement, also specifying how audio should complement the visual modality. **User testing** involves longitudinal tests, and detailed characterization of the participants including their audio-visual and motor impairments, in order to determine system **usability** and the impact of multimodality.

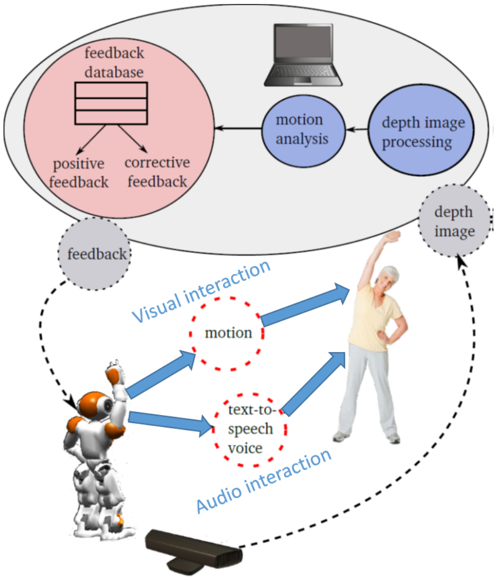


Figure 11 Step-by-step design of a multimodal interaction system for elderly people, namely a robotic fitness coach. Figure adapted from [Görer et al. 2016]. See text for additional details.

ability, perceived enjoyment, attitude, anxiety, social presence and influence, and trust, among other factors.

0.6 Conclusions

In this chapter we have surveyed a range of multimodal-multisensor technologies, and discussed how older adults' needs can be addressed through such technologies. Our survey establishes that although multimodal interaction has a lot of potential, present research effort tends to be concentrated on a small subset of these needs, including basic sustenance, social needs, independence, and health. Other needs such as personal growth, learning, and self-esteem are yet to be explored to their full potential.

With respect to specific application areas, we find that multimodal robotic assistants currently serve areas like health, fitness and social interaction, and their market is growing. Low-cost mobile technologies have significant potential for supporting older adults' access to information and for social connectedness. Personal assistants and smart homes represent significant opportunities for the development of elder-centred multimodal interfaces.

Assistive technologies for older adults have not seen the same level of integration of multimodal-multisensor interactions as other application areas. Most research has focused on combining speech and/or gestures with secondary modalities. The main drivers for multimodal research within gerontechnology have been increasing the usability, accessibility, and familiarity of interfaces, yet there are gaps between developments in multimodal interfaces and research in gerontechnology. Our research suggests balancing adaptivity and usability as design principles.

Researchers should be aware that this demographic group is characterized by a much wider range of abilities, capabilities, needs, and wants than other groups. Our analysis suggests that attention should be given to issues that may prevent the full adoption of multimodal interfaces by seniors. In particular, easy modality selection, integration of modalities sensitive to the physiological and cognitive challenges affecting older adults, and socio-cultural factors should be considered. This requires approaches solidly grounded in inclusive design principles, involving users in all stages of design, development, and evaluation.

Supplementary Materials

0.6.1 Focus questions

1. What are the use cases where multimodal interaction is most effective in supporting elderly users?
2. What potential issues can multimodal interfaces introduce for older adults? How can such issues be addressed? Discuss specific successful examples.
3. How do the elderly users differ from younger users? Design a user study for a tablet-based multimodal photo sharing application – indicate requirement gathering approaches and discuss usability evaluation measures and design guidelines.
4. What are the trends in research and commercialization for developing multimodal technologies for the elderly?
5. What are typical interface features and basic capabilities of systems that interact with elderly users?
6. What are the most important technology acceptance issues for the elderly, and how are these affected by cultural and societal factors?
7. What are the interface design cycle elements that are of particular importance when designing multimodal assistive technologies for seniors?
8. Which interface design considerations are most critical for elderly users, and why?
9. In what ways can interaction fail between a multimodal/multi-sensor system and an elderly user? What can designers do to prevent and mitigate these failures?
10. What are the most successful design approaches for developing technology for seniors, and why? Illustrate this by designing a wearable multimodal-multisensor interface for tracking seniors' activity.
11. Why is multimodal integration more difficult in elder-centred interfaces?
12. What are the modalities that are most preferred by seniors and why? What are the conditions under which these modalities are not suitable?

Bibliography

G. D. Abowd, A. F. Bobick, I. A. Essa, E. D. Mynatt, and W. A. Rogers. 2002. The aware home: A living laboratory for technologies for successful aging. In *Proc. AAAI-02 Workshop Automation as Caregiver*, pp. 1–7.

A. N. Aicha, S. Haitjema, and P. Wiggers. 2016. Hipperbox: A portable toolbox for rehabilitation from a hip surgery using sensors. In *Proceedings of Pervasive Health*.

Aldebaran Inc., 2016. Aldebaran Pepper. https://www.ald.softbankrobotics.com/en.

H. Alemdar and C. Ersoy. 2010. Wireless sensor networks for healthcare: A survey. *Computer Networks*, 54(15): 2688–2710.

Y. Aly and C. Munteanu. 2016. An information-centric framework for mobile collaboration between seniors and caregivers that balances independence, privacy, and social connectedness. In *International Conference on Human-Computer Interaction*, pp. 395–400. Springer.

Alzheimers' Association, 2016. Mild cognitive impairment. http://www.alz.org/what-is-dementia.asp.

Amazon Inc., 2016. Amazon Echo. http://www.amazon.com/oc/echo/.

ANSA, 2016. I primi robot-badanti. http://www.ansa.it/scienza/notizie/rubriche/tecnologie/2014/02/27/primi-robot-badanti_10155183.html.

Apple Inc., 2016. Apple Siri. http://www.apple.com/ios/siri.

A. Attarwala, C. Munteanu, and R. Baecker. 2013. An accessible, large-print, listening and talking e-book to support families reading together. In *Proc. 15th Int. Conf. on Human-computer Interaction with Mobile Devices and Services*, pp. 440–443. ACM, New York, NY, USA. DOI: 10.1145/2493190.2494658.

M. P. Aylett, P. O. Kristensson, S. Whittaker, and Y. Vazquez-Alvarez. 2014. None of a chind: relationship counselling for hci and speech technology. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pp. 749–760. ACM.

Y. Barnard, M. D. Bradley, F. Hodgson, and A. D. Lloyd. 2013. Learning to use new technologies by older adults: Perceived difficulties, experimentation behaviour and usability. *Computers in Human Behavior*, 29(4): 1715–1724.

P. Barnes, T. Spooner, J. Dhabolt, and J. Leighton. 2009. Instructional gaming methods and apparatus. *Patent A1 20090176581*.

M. Belshaw, B. Taati, J. Snoek, and A. Mihailidis. 2011. Towards a single sensor passive solution for automated fall detection. In *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 1773–1776. IEEE.

G. M. Bertolotti, A. M. Cristiani, P. Colagiorgio, F. Romano, E. Bassani, N. Caramia, and S. Ramat. 2016. A wearable and modular inertial unit for measuring limb movements and balance control abilities. *IEEE Sensors Journal*, 16(3): 790–797.

R. Bevilacqua, E. Felici, F. Marcellini, S. Glende, S. Klemcke, I. Conrad, R. Esposito, F. Cavallo, and P. Dario. 2015. Robot-era project: Preliminary results on the system usability. In *International Conference of Design, User Experience, and Usability*, pp. 553–561. Springer.

H. Beyer and K. Holtzblatt. 1997. *Contextual design: defining customer-centered systems*. Elsevier.

T. W. Bickmore, L. Caruso, and K. Clough-Gorr. 2005. Acceptance and usability of a relational agent interface by urban older adults. In *CHI*, pp. 1212–1215. ACM.

J. Bollwein, D. Volkert, R. Diekmann, M. J. Kaiser, W. Uter, K. Vidal, C. C. Sieber, and J. M. Bauer. 2013. Nutritional status according to the mini nutritional assessment (mna®) and frailty in community dwelling older persons: A close relationship. *The journal of nutrition, health & aging*, 17(4): 351–356. ISSN 1760-4788. http://dx.doi.org/10.1007/s12603-013-0034-7. DOI: 10.1007/s12603-013-0034-7.

I. Borges and D. Sinclair, December 2008. Media literacy, digital exclusion and older people. AGE Report - the European Union's Older People's Platform.

W. Bourennane, Y. Charlon, F. Bettahar, E. Campo, and D. Esteve. 2013. Homecare monitoring system: A technical proposal for the safety of the elderly experimented in an Alzheimer's care unit. *IRBM*, 34(2): 92–100.

P. Bovbel and G. Nejat. 2014. Casper: An assistive kitchen robot to promote aging in place. *Journal of Medical Devices*, 8(3): 030945.

British Columbia Ministry of Health, March 2004. Social isolation among seniors: an emerging issue. children's, women's and seniors' health branch report.

J. Brooke et al. 1996. SUS-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194): 4–7.

P. Busquin, 2013. Final evaluation of the ambient assisted living joint programme. http://www.aal-europe.eu/wp-content/uploads/2015/05/Final-report-of-the-AAL-Busquin-2013.pdf.

S. P. Carmien and G. Fischer. 2008. Design, adoption, and assessment of a socio-technical environment supporting independence for persons with cognitive disabilities. In *CHI*, pp. 597–606. ACM, New York, NY, USA. ISBN 978-1-60558-011-1. DOI: 10.1145/1357054.1357151.

J. C. Castillo, D. Carneiro, J. Serrano-Cuerda, P. Novais, A. Fernández-Caballero, and J. Neves. 2014. A multi-modal approach for activity classification and fall detection. *International Journal of Systems Science*, 45(4): 810–824.

M. Chan, D. Estève, C. Escriba, and E. Campo. 2008. A review of smart homes—present state and future challenges. *Computer methods and programs in biomedicine*, 91(1): 55–81.

H.-E. Chen, Y.-Y. Lin, C.-H. Chen, I. Wang, et al. 2015. Blindnavi: a navigation app for the visually impaired smartphone user. In *Proc. 33rd Annual ACM Conf. Extended Abstracts on Human Factors in Computing Systems – CHI*, pp. 19–24. ACM.

Y.-R. R. Chen and P. J. Schulz. 2016. The effect of information communication technology interventions on reducing social isolation in the elderly: A systematic review. *Journal of medical Internet research*, 18(1).

J. Coelho and C. Duarte. 2011. The contribution of multimodal adaptation techniques to the guide interface. In *International Conference on Universal Access in Human-Computer Interaction*, pp. 337–346. Springer.

J. Coelho, C. Duarte, P. Biswas, and P. Langdon. 2011. Developing accessible tv applications. In *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, pp. 131–138. ACM.

P. Cohen, A. Cheyer, E. Horvitz, R. El Kaliouby, and S. Whittaker. 2016. On the future of personal assistants. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 1032–1037. ACM.

R. Colombo, F. Pisano, A. Mazzone, C. Delconte, S. Micera, M. Carrozza, P. Dario, and G. Minuco. 2007. Design strategies to improve patient motivation during robot-aided rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 4(1): 3–12.

Conference Board of Canada, 2010. Adult literacy rate – low level skills. education report.

S. J. Czaja and C. C. Lee. 2002. Designing computer systems for older adults. In *The human-computer interaction handbook*, pp. 413–427. L. Erlbaum Associates Inc.

S. Czarnuch, A. Mihailidis, et al. 2011. The design of intelligent in-home assistive technologies: Assessing the needs of older adults with dementia and their caregivers. *Gerontechnology*, 10(3): 169–182.

P. Dadlani, A. Sinitsyn, W. Fontijn, and P. Markopoulos. 2010. Aurama: Caregiver awareness for living independently with an augmented picture frame display. *AI Soc.*, 25(2): 233–245. ISSN 0951-5666. DOI: 10.1007/s00146-009-0253-y.

G. Demiris and B. K. Hensel. 2008. Technologies for an aging society: a systematic review of "smart home" applications. *Yearbook of Medical Informatics*, 3: 33–40.

B. E. Demiröz, A. A. Salah, and L. Akarun. 2014. Coupling fall detection and tracking in omnidirectional cameras. In *Int. Workshop on Human Behavior Understanding*, pp. 73–85. Springer.

D. DeVault, R. Artstein, G. Benn, T. Dey, E. Fast, A. Gainer, K. Georgila, J. Gratch, A. Hartholt, M. Lhommet, et al. 2014. Simsensei kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pp. 1061–1068. International Foundation for Autonomous Agents and Multiagent Systems.

E. Dishman. 2004. Inventing wellness systems for aging in place. *Computer*, 37(5): 34–41.

B. A. Farshchian and Y. Dahl. 2015. The role of ict in addressing the challenges of age-related falls: a research agenda based on a systematic mapping of the literature. *Personal and Ubiquitous Computing*, 19(3-4): 649–666.

J. Fasola and M. Mataric. Aug 2012. Using socially assistive human-robot interaction to motivate physical exercise for older adults. *Proceedings of the IEEE*, 100(8): 2512–2526.

J. Fasola and M. Mataric. 2013. A socially assistive robot exercise coach for the elderly. *Journal of Human-Robot Interaction*, 2(2): 3–32.

D. Feil-Seifer and M. Mataric. 2005. Defining socially assistive robotics. In *9th International Conference on Rehabilitation Robotics, ICORR*, pp. 465–468.

F. Ferreira, N. Almeida, A. F. Rosa, A. Oliveira, A. Teixeira, and J. C. Pereira. 2013. Multimodal and adaptable medication assistant for the elderly: A prototype for interaction and usability in smartphones. In *8th Iberian Conf. on Information Systems and Technologies*, pp. 1–6. IEEE.

F. Ferreira, N. Almeida, A. F. Rosa, A. Oliveira, J. Casimiro, S. Silva, and A. Teixeira. 2014. Elderly centered design for interaction–the case of the s4s medication assistant. *Procedia Computer Science*, 27: 398–408.

S. Follmer, R. T. Ballagas, H. Raffle, M. Spasojevic, and H. Ishii. 2012. People in books: using a flashcam to become part of an interactive book for connected reading. In *Proceedings of the ACM 2012 conference on Computer supported cooperative work*, pp. 685–694. ACM.

R. L. Franz, C. Munteanu, B. B. Neves, and R. Baecker. 2015. Time to retire old methodologies? reflecting on conducting usability evaluations with older adults. In *Proc. 17th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services Adjunct*, pp. 912–915. ACM.

E. Freeman, G. Wilson, D. Vo, A. Ng, I. Politis, and S. Brewster. 2017X. Multimodal feedback in HCI: Haptics, non-speech audio, and their applications. In S. Oviatt, B. Schuller, P. Cohen, D. Sonntag, G. Potamianos, and A. Krueger, eds., *The Handbook of Multimodal-Multisensor Interfaces: Language Processing, Software, Commercialization, and Emerging Directions*, pp. xxx–xxx. Morgan Claypool Publishers, San Rafael, CA, USA.

GiraffPlus, 2016. The GiraffPlus Project. http://giraffplus.eu/.

B. Görer, A. A. Salah, and H. L. Akın. 2013. A robotic fitness coach for the elderly. In *Ambient Intelligence*, pp. 124–139. Springer.

B. Görer, A. A. Salah, and H. L. Akın. 2016. An autonomous robotic exercise tutor for elderly people. *Autonomous Robots*. DOI: 10.1007/s10514-016-9598-5.

J. Gratch, L.-P. Morency, S. Scherer, G. Stratou, J. Boberg, S. Koenig, T. Adamson, A. A. Rizzo, et al. 2013. User-state sensing for virtual health agents and telehealth applications. In *MMVR*, pp. 151–157.

A. Hämäläinen, A. Teixeira, N. Almeida, H. Meinedo, T. Fegyó, and M. S. Dias. 2015. Multilingual speech recognition for the elderly: The aalfred personal life assistant. *Procedia Computer Science*, 67: 283–292.

T. Han, X. Xiao, L. Shi, J. Canny, and J. Wang. 2015. Balancing accuracy and fun: Designing camera based mobile games for implicit heart rate monitoring. In *CHI*, pp. 847–856. ACM, New York, NY, USA. ISBN 978-1-4503-3145-6. DOI: 10.1145/2702123.2702502.

C. M. Harrison. 2004. Low-vision reading aids: reading as a pleasurable experience. *Personal and Ubiquitous Computing*, 8(3-4): 213–220.

H. Hawley-Hague, E. Boulton, A. Hall, K. Pfeiffer, and C. Todd. 2014. Older adults' perceptions of technologies aimed at falls prevention, detection or monitoring: a systematic review. *International journal of medical informatics*, 83(6): 416–426.

M. Heerink. 2010. *Assessing acceptance of assistive social robots by aging adults*. PhD thesis, University of Amsterdam.

M. Heerink, B. Kröse, V. Evers, and B. Wielinga. 2010. Assessing acceptance of assistive social agent technology by older adults: The Almere model. *International Journal of Social Robotics*, 2(4): 361–375.

G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, et al. 2012. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6): 82–97.

X. Huang, J. Baker, and R. Reddy. 2014. A historical perspective of speech recognition. *Communications of the ACM*, 57(1): 94–103.

J. A. Jacko, I. U. Scott, F. Sainfort, K. P. Moloney, T. Kongnakorn, B. S. Zorich, and V. K. Emery. 2002. Effects of multimodal feedback on the performance of older adults with normal and impaired vision. In *Universal Access Theoretical Perspectives, Practice, and Experience*, pp. 3–22. Springer.

C. Jian, H. Shi, N. Sasse, C. Rachuy, F. Schafmeister, H. Schmidt, and N. von Steinbüchel. 2014. Modality preference in multimodal interaction for elderly persons. In *Biomedical Engineering Systems and Technologies*, pp. 378–393. Springer.

Jibo Inc., 2016. Jibo, the world's first social robot. https://www.jibo.com/.

A. Kalogianni, 2015. Toyota jumpstarts robotic elderly care with the HSR robot prototype. http://www.digitaltrends.com/cars/toyota-develops-human-support-robot-for-elder-care.

P. Kannus, H. Sievänen, M. Palvanen, T. Järvinen, and J. Parkkari. 2005. Prevention of falls and consequent injuries in elderly people. *The Lancet*, 366(9500): 1885–1893.

S. Kemper and T. L. Mitzner. 2001. Language production and comprehension. *Handbook of the psychology of aging*, 5: 378–398.

S. Kieffer, J.-Y. L. Lawson, and B. Macq. 2009. User-centered design and fast prototyping of an ambient assisted living system for elderly people. In *Information Technology: New Generations, 2009. ITNG'09. Sixth International Conference on*, pp. 1220–1225. IEEE.

S. Kumar, W. J. Nilsen, A. Abernethy, A. Atienza, K. Patrick, M. Pavel, W. T. Riley, A. Shar, B. Spring, D. Spruijt-Metz, et al. 2013. Mobile health technology evaluation: the mhealth evidence workshop. *American journal of preventive medicine*, 45(2): 228–236.

G. Lepicard and N. Vigouroux. 2010. Touch screen user interfaces for older subjects. In *Computers Helping People with Special Needs*, pp. 592–599. Springer.

R. Looije, M. A. Neerincx, and F. Cnossen. 2010. Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. *International Journal of Human-Computer Studies*, 68(6): 386–397.

D. Lopez Recio, L. Marquez Segura, E. Marquez Segura, and A. Waern. 2013. The NAO models for the elderly. In *8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 187–188.

K. MacLean, O. Schneider, and H. Seifi. 2017X. Multisensory haptic interactions: Understanding the sense and designing for it. In S. Oviatt, B. Schuller, P. Cohen, D. Sonntag, G. Potamianos, and A. Krueger, eds., *The Handbook of Multimodal-Multisensor Interfaces: Language Processing, Software, Commercialization, and Emerging Directions*, pp. xxx–xxx. Morgan Claypool Publishers, San Rafael, CA, USA.

M. H. Mallers, M. Claver, and L. A. Lares. 2014. Perceived control in the lives of older adults: The influence of langer and rodin's work on gerontological theory, policy, and practice. *The Gerontologist*, 54(1): 67–74.

A. H. Maslow. 1943. A theory of human motivation. *Psychological review*, 50(4): 370.

A. H. Maslow. 1954. *Motivation and personality*. Harper and Row.

D. McColl and G. Nejat. 2014. Recognizing emotional body language displayed by a human-like social robot. *International Journal of Social Robotics*, 6(2): 261–280.

D. McColl, W.-Y. G. Louie, and G. Nejat. 2013. Brian 2.1: A socially assistive robot for the elderly and cognitively impaired. *IEEE Robotics & Automation Magazine*, 20(1): 74–83.

M. R. McGee-Lennon, M. K. Wolters, and S. Brewster. 2011. User-centred multimodal reminders for assistive living. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2105–2114. ACM.

R. K. Megalingam, A. Nambissan, A. Thambi, A. Gopinath, and K. Megha. 2015. Sound-and touch-based smart cane: Better walking experience for visually challenged. In *Intelligent Computing, Communication and Devices*, pp. 589–595. Springer.

S. Mellone, C. Tacconi, L. Schwickert, J. Klenk, C. Becker, and L. Chiari. 2012. Smartphone-based solutions for fall detection and prevention: the farseeing approach. *Zeitschrift für Gerontologie und Geriatrie*, 45(8): 722–727.

M. Memon, S. R. Wagner, C. F. Pedersen, F. H. A. Beevi, and F. O. Hansen. 2014. Ambient assisted living healthcare frameworks, platforms, standards, and quality attributes. *Sensors*, 14(3): 4312–4341.

K. Moffatt. 2013. Older-adult hci: Why should we care? *interactions*, 20(4): 72–75.

C. Mollaret, A. Mekonnen, F. Lerasle, I. Ferrané, J. Pinquier, B. Boudet, and P. Rumeau. 2016. A multi-modal perception based assistive robotic system for the elderly. *Computer Vision and Image Understanding*.

L.-P. Morency, G. Stratou, D. DeVault, A. Hartholt, M. Lhommet, G. M. Lucas, F. Morbini, K. Georgila, S. Scherer, J. Gratch, et al. 2015. SimSensei demonstration: A perceptive virtual human interviewer for healthcare applications. In *AAAI*, pp. 4307–4308.

D. Muñoz, F. J. Gutierrez, and S. F. Ochoa. 2015. Introducing ambient assisted living technology at the home of the elderly: Challenges and lessons learned. In *International Workshop on Ambient Assisted Living*, pp. 125–136. Springer.

C. Munteanu, H. Molyneaux, D. McDonald, J. Lumsden, R. Leung, H. Fournier, and J. Maitland. 2011. “showing off” your mobile device: Adult literacy learning in the classroom and beyond. *Proc. of MobileHCI 2011*.

C. Munteanu, H. Molyneaux, J. Maitland, D. McDonald, R. Leung, H. Fournier, and J. Lumsden. 2014. Hidden in plain sight: low-literacy adults in a developed country overcoming social and educational challenges through mobile learning support tools. *Personal and ubiquitous computing*, 18(6): 1455–1469.

A. B. Naumann, I. Wechsung, and J. Hurtienne. 2010. Multimodal interaction: A suitable strategy for including older users? *Interacting with Computers*, 22(6): 465–474.

C. Neustaedter and S. Greenberg. 2012. Intimacy in long-distance relationships over video chat. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 753–762. ACM.

B. B. Neves, F. Amaro, and J. R. Fonseca. 2013. Coming of (old) age in the digital age: ICT usage and non-usage among older adults. *Sociological Research Online*, 18(2): 6.

B. B. Neves, R. L. Franz, C. Munteanu, R. Baecker, and M. Ngo. 2015. “my hand doesn’t listen to me!”: Adoption and evaluation of a communication technology for the ‘oldest old’. In *CHI*, pp. 1593–1602. ACM, New York, NY, USA. ISBN 978-1-4503-3145-6. DOI: 10.1145/2702123.2702430.

N. R. Nicholson. 2012. A review of social isolation: an important but underassessed condition in older adults. *The journal of primary prevention*, 33(2-3): 137–152.

E. Nicol, A. Komninos, and M. D. Dunlop. 2016. A participatory design and formal study investigation into mobile text entry for older adults. *International Journal of Mobile Human Computer Interaction*, 8(2): 20–46.

M. Nilsson, J. Ingvast, J. Wikander, and H. von Holst. 2012. The soft extra muscle system for improving the grasping capability in neurological rehabilitation. In *Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on*, pp. 412–417. IEEE.

K. L. Norman. 2013. GEQ (game engagement/experience questionnaire): a review of two papers. *Interacting with Computers*, 25(4): 278–283.

E. O'Brien, 2016. Older adults buddy up with Amazon's Alexa. MarketWatch, 18 March 2016. http://www.marketwatch.com/story/older-adults-buddy-up-with-amazons-alexa-2016-03-18. Retrieved on 10 August 2016.

S. Oviatt. 2002. Breaking the robustness barrier: Recent progress on the design of robust multimodal systems. In M. Zelkowitz, ed., *Advances in Computers*, volume 56, pp. 305–341. Academic Press.

S. Oviatt. 2003. Advances in robust multimodal interface design. *IEEE Computer Graphics and Applications*, 23(5): 62–68.

S. Oviatt, R. Coulston, S. Tomko, B. Xiao, R. Lunsford, M. Wesson, and L. Carmichael. 2003. Toward a theory of organized multimodal integration patterns during human-computer interaction. In *Proceedings of the 5th international conference on Multimodal interfaces*, pp. 44–51. ACM.

Paro Robots U.S. Inc., 2016. The PARO therapeutic robot. http://www.parorobots.com/.

A. Passerini and M. Sebag. 2015. Learning and optimization with the human in the loop. *Constraints, Optimization and Data*, p. 21.

S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers. 2012. A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation*, 9(1): 1.

C. M. Perissinotto, I. S. Cenzer, and K. E. Covinsky. 2012. Loneliness in older persons: a predictor of functional decline and death. *Archives of internal medicine*, 172(14): 1078–1084.

M. Perry, A. Dowdall, L. Lines, and K. Hone. 2004. Multimodal and ubiquitous computing systems: supporting independent-living older users. *Information Technology in Biomedicine, IEEE Transactions on*, 8(3): 258–270.

B. Pfleging, M. Kienast, A. Schmidt, T. Döring, et al. 2011. Speet: A multimodal interaction style combining speech and touch interaction in automotive environments. In *Adjunct Proc. 3rd Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, volume 11.

J. Pineau, M. Montemerlo, M. Pollack, N. Roy, and S. Thrun. 2003. Towards robotic assistants in nursing homes: Challenges and results. *Robotics and Autonomous Systems*, 42(3): 271–281.

C. G. Pires, F. Pinto, V. D. Teixeira, J. Freitas, and M. S. Dias. 2012. Living home center–a personal assistant with multimodal interaction for elderly and mobility impaired e-inclusion. *Proc. PROPOR. Coimbra*.

A. Prange, I. P. Sandrala, M. Weber, and D. Sonntag. 2015. Robot companions and smartpens for improved social communication of dementia patients. In *Proceedings of the 20th International Conference on Intelligent User Interfaces Companion*, pp. 65–68. ACM.

J. Preece, Y. Rogers, and H. Sharp. 2015. *Interaction design: beyond human-computer interaction*, 4th. Wiley.

M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng. 2009. Ros: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, volume 3, p. 5.

P. Rashidi and A. Mihailidis. May 2013. A survey on ambient-assisted living tools for older adults. *Biomedical and Health Informatics, IEEE Journal of*, 17(3): 579–590. ISSN 2168-2194. DOI: 10.1109/JBHI.2012.2234129.

K. Rector, C. L. Bennett, and J. A. Kientz. 2013. Eyes-free yoga: An exergame using depth cameras for blind & low vision exercise. In *Proc. 15th Int. ACM SIGACCESS Conf. on Computers and Accessibility*, pp. 12:1–12:8. ACM, New York, NY, USA. ISBN 978-1-4503-2405-2. DOI: 10.1145/2513383.2513392.

Y. Rogers. 2011. *Interaction design : beyond human-computer interaction*. Wiley. ISBN 9780470665763.

A. A. Salah, B. J. Kröse, and D. J. Cook. 2015. Behavior analysis for elderly. In *Proc. 6th Int. Workshop on Human Behavior Understanding*, pp. 1–10. Springer.

T. A. Salthouse. 2016. *Theoretical perspectives on cognitive aging*. Routledge.

T. Sato and J. E. Cameron. 1999. The relationship between collective self-esteem and self-construal in japan and canada. *The Journal of Social Psychology*, 139(4): 426–435.

D. Schnelle-Walka and S. Radomski. 2017X. Automotive multimodal human-machine interaction. In S. Oviatt, B. Schuller, P. Cohen, D. Sonntag, G. Potamianos, and A. Krueger, eds., *The Handbook of Multimodal-Multisensor Interfaces: Language Processing, Software, Commercialization, and Emerging Directions*, pp. xxx–xxx. Morgan Claypool Publishers, San Rafael, CA, USA.

D. Schuler and A. Namioka. 1993. *Participatory design: Principles and practices*. CRC Press.

R. Shilkrot, J. Huber, W. Meng Ee, P. Maes, and S. C. Nanayakkara. 2015. Fingerreader: a wearable device to explore printed text on the go. In *CHI*, pp. 2363–2372. ACM.

E. Siegenthaler, P. Wurtz, and R. Groner. 2010. Improving the usability of e-book readers. *Journal of Usability Studies*, 6(1): 25–38.

K. A. Siek, Y. Rogers, and K. H. Connelly. 2005. Fat finger worries: how older and younger users physically interact with pdas. In *IFIP Conference on Human-Computer Interaction*, pp. 267–280. Springer.

D. Y. Y. Sim and C. K. Loo. 2015. Extensive assessment and evaluation methodologies on assistive social robots for modelling human–robot interaction–a review. *Information Sciences*, 301: 305–344.

S. Solaimani, W. Keijzer-Broers, and H. Bouwman. 2013. What we do–and don't–know about the smart home: an analysis of the smart home literature. *Indoor and Built Environment*, p. 1420326X13516350.

R. Stiefelhagen, H. K. Ekenel, C. Fugen, P. Gieselmann, H. Holzapfel, F. Kraft, K. Nickel, M. Voit, and A. Waibel. 2007. Enabling multimodal human–robot interaction for the karlsruhe humanoid robot. *IEEE Transactions on Robotics*, 23(5): 840–851.

L. Sucar, G. Azcárate, R. Leder, D. Reinkensmeyer, J. Hernández, I. Sanchez, and P. Saucedo. 2009. Gesture therapy: A vision-based system for arm rehabilitation after stroke. In *Biomedical Engineering Systems and Technologies International Joint Conference, BIOSTEC*, volume 25, pp. 531–540. Springer.

K. Tanaka, S. Mu, and S. Nakashima. 2014. Meal-assistance robot using ultrasonic motor with eye interface. *Journal ref: International Journal of Automation Technology*, 8(2): 186–192.

A. Tapus and M. J. Mataric. 2008. Socially assistive robots: The link between personality, empathy, physiological signals, and task performance. In *AAAI Spring Symposium: Emotion, Personality, and Social Behavior*, pp. 133–140.

A. Teixeira, A. Hämäläinen, J. Avelar, N. Almeida, G. Németh, T. Fegyó, C. Zainkó, T. Csapó, B. Tóth, A. Oliveira, et al. 2014. Speech-centric multimodal interaction for easy-to-access online services–a personal life assistant for the elderly. *Procedia computer science*, 27: 389–397.

V. Teixeira, C. Pires, F. Pinto, J. Freitas, M. S. Dias, and E. M. Rodrigues. 2012. Towards elderly social integration using a multimodal human-computer interface. In *Proc. International Living Usability Lab Workshop on AAL Latest Solutions, Trends and Applications, AAL*.

S. Thielke, M. Harniss, H. Thompson, S. Patel, G. Demiris, and K. Johnson. 2011. Maslow's hierarchy of human needs and the adoption of health-related technologies for older adults. *Ageing International*, 37(4): 470–488. ISSN 1936-606X. http://dx.doi.org/10.1007/s12126-011-9121-4. DOI: 10.1007/s12126-011-9121-4.

A. K. Thomas, J. B. Dave, and B. M. Bonura. 2010. Theoretical perspectives on cognitive aging. In *Handbook of Medical Neuropsychology*, pp. 297–313. Springer.

M. E. Tinetti, S. T. Bogardus Jr, J. V. Agostini, et al. 2004. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. *N Engl J Med*, 351(27): 2870–2874.

M. Turk. 2014. Multimodal interaction: A review. *Pattern Recogn. Lett.*, 36: 189–195.

S. Uzor, L. Baillie, and D. Skelton. 2012. Senior designers: empowering seniors to design enjoyable falls rehabilitation tools. In *CHI*, pp. 1179–1188. ACM.

A. van Berlo. 2002. Smart home technology: Have older people paved the way? *Gerontechnology*, 2(1): 77–87.

V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS Quarterly*, pp. 425–478.

J. Vines, M. Blythe, S. Lindsay, P. Dunphy, A. Monk, and P. Olivier. 2012. Questionable concepts: critique as resource for designing with eighty somethings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1169–1178. ACM.

R. Vipperla, S. Renals, and J. Frankel. 2008. Longitudinal study of asr performance on ageing voices.

T. Visser, M. Vastenburg, and D. Keyson. 2010. Snowglobe: the development of a prototype awareness system for longitudinal field studies. In *Proc. 8th ACM Conference on Designing Interactive Systems*, pp. 426–429. ACM.

K. Wada, T. Shibata, T. Saito, K. Sakamoto, and K. Tanie. 2005. Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation,ICRA 2005*, pp. 2785–2790.

J. Waycott, H. Davis, F. Vetere, A. Morgans, A. Gruner, E. Ozanne, and L. Kulik. 2014. Captioned photographs in psychosocial aged care: relationship building and boundary work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 4167–4176. ACM.

Willow Garage Inc., 2016. Willow Garage's PR2 robot for research and innovation. http://www.willowgarage.com/pages/pr2/overview.

D. Wu, L. Pigou, P.-J. Kindermans, L. Nam, L. Shao, J. Dambre, and J.-M. Odobez. 2016. Deep dynamic neural networks for multimodal gesture segmentation and recognition.

B. Xiao, R. Lunsford, R. Coulston, M. Wesson, and S. Oviatt. 2003. Modeling multimodal integration patterns and performance in seniors: toward adaptive processing of individual differences. In *Proc. 5th International Conference on Multimodal Interfaces*, pp. 265–272. ACM.

R. Yaghoubzadeh, M. Kramer, K. Pitsch, and S. Kopp. 2013. Virtual agents as daily assistants for elderly or cognitively impaired people. In *International Workshop on Intelligent Virtual Agents*, pp. 79–91. Springer.

L. Yang, H. Cheng, J. Hao, Y. Ji, and Y. Kuang. 2015. A survey on media interaction in social robotics. In *Advances in Multimedia Information Processing–PCM 2015*, pp. 181–190. Springer.

G. Yavuz, M. Kocak, G. Ergun, H. O. Alemdar, H. Yalcin, O. D. Incel, and C. Ersoy. 2010. A smartphone based fall detector with online location support. In *Proceedings of the International Workshop on Sensing for App Phones, Zurich, Switzerland*, pp. 31–35.

M. Zajicek and W. Morrissey. 2003. Multimodality and interactional differences in older adults. *Universal Access in the Information Society*, 2(2): 125–133.

M. Ziefle, P. Pappachan, E.-M. Jakobs, and H. Wallentowitz. 2008. Visual and auditory interfaces of advanced driver assistant systems for older drivers. In *Proc. ICCHP, Linz, Austria*. Springer Verlag.

S. A. Zwijsen, A. R. Niemeijer, and C. M. Hertogh. 2011. Ethics of using assistive technology in the care for community-dwelling elderly people: An overview of the literature. *Aging & mental health*, 15(4): 419–427.