Universal Access in the Information Society
https://doi.org/10.1007/s10209-023-01037-3

LONG PAPER

An Interaction Framework for Designing Systems for Virtual Home Assistants and People with Dysarthria

Aisha Jaddoh[1,2] · Fernando Loizides[1] · Jimin Lee[3] · Omer Rana[1]

Accepted: 17 August 2023

Abstract
People with speech impairments usually use assistive technology devices to assist them with communication and daily tasks. These devices can be controlled using different modalities, such as touch, eye gaze, gestures, and others. This article proposes a standardized methodology for designing non-verbal voice cue interactive systems that enable people with dysarthria to vocally interact with virtual home assistants (VHAs). We adopted a qualitative data-gathering approach to gain insights into users' experiences and requirements and to determine crucial design elements for designing interactive voice assistants for people with dysarthria. Nineteen participants with varying levels of dysarthria took part in the study to create the framework. A system was then built using the proposed framework, and an additional test was performed with a further seven participants to validate the created system, thus inferring the validity of the framework. Our work empirically demonstrates how an informed, structured design of a fast, direct (verbal rather than forcing users to change modalities or using an intermediate device) method of communication improves the usability of VHAs for people with dysarthria while simultaneously allowing for a more authentic experience. The current data also highlighted that using non-verbal voice cues would be a convenient option. By providing a reproducible framework for developing non-verbal interactive systems for VHAs, we can increase the accessibility of said devices, their usability, and their user experience.

Keywords Non-verbal voice cues · Framework design · Dysarthria · Virtual home assistants

1 Introduction

In their daily lives, people with disabilities rely on various assistive devices or systems to communicate with others, ensure mobility, and perform daily tasks. These devices include physical assistive aids, such as a wheelchair for a physical disability, a white cane for blindness, or an augmentative and alternative communication device for cognitive or speech disabilities. Further, they may use software and computers for screen reader applications and voice synthesis speech systems. In addition to using these devices, they have begun to use voice interfaces, such as Apple's Siri and Microsoft's Cortana. Similarly, standalone voice interface devices, or virtual home assistants (VHAs), such as Google Home and Amazon's Alexa, can help people with disabilities to perform their daily tasks. VHAs are devices that receive voice commands and perform tasks accordingly. For example, a person with vision impairment could use a VHA to control other devices or appliances without assistance from another person or device. Moreover, a person with a physical disability could use a VHA to perform a task that requires moving without having to perform any physical movement.

However, a question that arises is whether these devices are fully accessible. Studies have recognized that some people with disabilities, such as persons with speech impairment, may not be able to use VHAs. For example, Glasser [1] showed that individuals who are deaf or hard of hearing

✉ Aisha Jaddoh
jaddoha@cardiff.ac.uk

Fernando Loizides
loizidesf@cardiff.ac.uk

Jimin Lee
jxl91@psu.edu

Omer Rana
ranaof@cardiff.ac.uk

[1] School of Computer Science and Informatics, Cardiff University, Cardiff, UK

[2] Department of Computer Science and Engineering, Yanbu Industrial College, Yanbu, Saudi Arabia

[3] Communication Sciences and Disorders Department, The Pennsylvania State University, State College, USA

Published online: 05 September 2023

and have speech difficulties face challenges when interacting with VHAs. Similarly, Ballati [2] demonstrated that VHAs do not perform at a sufficient level for people with speech impairment the more severe the case, the worse the performance.

In this study, we focused on one type of speech impairment: dysarthria. Dysarthria is a neurological motor speech disorder; it is one of the most severe communication disorders [3].The speech of an individual with dysarthria is characterized by slowness and a low level of intelligibility, and such individuals find it difficult to control the movement of their speech muscles [4].

Although people with speech impairment can use different input forms, such as a mouse and a keyboard, dysarthria can co-occur with other physical disabilities [5],which limits the use of those forms of input. This aspect restricts their independence as regards communicating as well as performing daily tasks. This is the case for some who have experienced traumatic brain injuries and people with cerebral palsy or some progressive neuralgic impairments. Therefore, it is worthwhile to explore the use of voice as method of interaction [6].

In this context, some studies have examined the challenges that individuals with dysarthria face in using VHAs, but have not proposed potential solutions, whereas other studies have proposed solutions. For instance, in one study, brain signals were converted to speech commands that were sent to a VHA [7]. However, this type of solution is not yet sufficiently advanced for commercial or daily use by people with dysarthria.

To fill this literature gap, in this article, we describe a technique that relies solely on speech, that is, non-verbal voice cues (e.g., “ahh”), which enable people with dysarthria to interact directly with a device without using an intermediary device or system to type or send commands. This approach would be especially useful for people with both dysarthria and a physical disability. We also describe the inclusive design method we used to create a system that aims to improve usability.

We conducted interviews with people with dysarthria in order to assess their needs and requirements. The focus of the interviews was to understand how they use technology for assistance with their daily tasks. In addition to asking them about their experience using voice assistants in general or VHAs in particular, we also asked them about their views on how systems should be designed to use non-verbal voice cues and account for their enunciation capabilities and preferences. Data collected from these interviews contributed to the system design.

1.1 Related work

Numerous studies have been conducted on speech technologies and people with dysarthria, and in particular, on automatic speech recognition (ASR) systems, and on different aspects of this domain. For example, in a review, Mustafa [8], explored how speech intelligibility affects the accuracy of ASR systems when used by people with dysarthria. Speech intelligibility has been defined as the level to which the listener can understand the speech of people with dysarthria [9]; Researchers have found it to be a factor that affects ASR accuracy. A direct correlation was found, which means the higher (lower) the intelligibility, the better (lower) the accuracy [10–14].

Many literature reviews have also focused on dysarthria and the ASR system. For instance, Rosen [13] identified different types of ASR systems and their performance when used by people with dysarthria. Next, Young [14] demonstrated challenges that people with dysarthria likely face when using such a system, such as the need to train the system, user fatigue, and errors that cause frustration. If systems require training, it is necessary to have a large dataset that contains examples of dysarthric speech. However, this might be difficult since collecting such recordings may cause fatigue [5, 15], which means recording sessions need to be longer in duration and include breaks or multiple short sessions [11].

Further, studies have focused on applications that can be implemented to aid people with dysarthria in overcoming the challenges related to ASR [16]. An early example is the Speech Training and Recognition for Dysarthric Users of Speech Technology project [17], which aims to help people with dysarthria to interact with assistive technologies. One application developed from this project is the voice-input voice-output communication aid [18], which converts impaired speech to synthetic speech. The user can create messages (simple or complex) from a small set of words. Another application from this project is that developed by Green [19] namely, an ASR recognizer for severe dysarthric speech. Both applications use isolated words as inputs. Another example of speech technology is CanSpeak [20], a customizable speech interface. Simple-to-pronounce keywords were mapped with litter, number, and commands to perform certain tasks. The system does not require training, for it is customized to each user. In contrast, ALADIN, another application [21, 22], requires training, but the training should be performed by the user, which means that the system fully adapts to the user’s speech. Similarly, Kim et al. [23] developed a voice user-interface keyboard for people with dysarthria that is customized to the user. The user needs to select a list of customized words: if they find it difficult to pronounce a particular word or the system does not understand it, they can pronounce another favorite word.

1.1.1 Virtual home assistants

In contrast to ASR systems, virtual home assistants have not been studied extensively by researchers. Thus, in this section we will also include research about different voice assistants, such as Siri. One study is by Ballati [24], who argues that the accessibility to virtual assistants needs improvement. Their study shows that virtual assistants can recognize between 47%-59% of dysarthric speech. This result was after evaluating the behavior of three virtual assistants (Siri, Amazon Alexa, and Google Assistant). In this domain, not only speech transcription accuracy affects the system's performance, but also the context of the words spoken. Moreover, the system's performance could be leveraged by a word that was recognized [24], using which the system can respond correctly. Similar to Ballati, Russis [25] also evaluated voice platforms. What is interesting in the latter's results is that the rate of correct transcription for the speakers was one sentence or less from the 51 sentences for all platforms. Moreover, the word error rate (WER) was high with a percentage of error higher than 78. While the research did not consider that the platform could yield results based on other factors, such as context, and not just transcribed text, the results clearly show that there is a limitation. Another research by Ballati [2] studied the accuracy of Siri, Google Assistant, and Cortana for the Italian language and concluded that the behavior of the voice assistants varies according to the user.

The dearth of studies in this domain may affect the validity of the results as mentioned by others, that is, using recorded sentences for databases. For example, the TORGO database [26] was not directly designed for the voice assistant, which made the sentences inadequate for use as commands for voice assistance. However, other studies recorded datasets specifically for this type of system.

1.1.2 Non-verbal interaction

Speech, as an input, is faster than using a mouse or a keyboard, especially for people with dysarthria. Nevertheless, as speech may be distorted, some researchers have studied alternatives. For instance, one vowel sound, /a/, was used in a keyword-spotting model, which triggers wake-up words [27] When interacting with humans, the use of sentences and continuous speech is closer to normal speech. However, non-verbal voice inputs also have advantages; because dysarthric speech has a slower rate of expression, non-verbal voice interactions may allow users to interact more quickly with the system.

Further, Spokra [28], developed a system, namely CHANTI for people with physical disabilities and speech impairments, which supported the use of different voice tones to control keyboard inputs. This input method has been tested in other contexts as well. Various studies have used non-verbal voice as a technique to control applications. Early examples include a study by Igarashi and Hughes [29], who used vowel sounds to control an interactive application in which the user finds a command and the application responds. Other studies have also used vowel sounds as inputs, including the use of a "vocal joystick" [30] as well as whistling and humming, to control a mouse cursor. In another context, Harada et al. [31] attempted to improve the accessibility of computer games by building a prototype that responded to non-verbal voice input.

1.2 Purpose of the Study

The purpose of this study was to address the challenges faced by people with dysarthria when using VHAs, through the two studies described in this article. The first study used data from interviews to explore the VHA experiences of individuals with dysarthria, thus improving the understanding of the issues they face and areas that require improvement. From the interview data, we collected user feedback and requirements regarding the proposed non-verbal voice cue interaction system. The results of this initial study determined the second phase, which is, designing a system tailored to user requirements. In the design process, we implemented a prototype to conduct preliminary testing of the proposed system. In summary, this study contributes to this area of research by exploring the accessibility to, and usability of, VHAs by individuals with dysarthria.

2 First Study: Interviews

2.1 Method

In this study, we aimed to design a non-verbal voice cue enabled system that is accessible and can be used by those with dysarthria and speech impairments. To this end, we needed a holistic set of requirements from representative users and therefore, we used both a participatory design method and a user-centered design approach. We adopted the interview questions from [23] and [32]. Both studies targeted the same audience as we do, and their general goal intersects with ours. Each of the studies examined the problem from a different angle. One focused on the user experience [23] and the other [32] on system design for people with dysarthria. Then, we used the responses from the interviews to design a new system.

Table 1 Participants' Age Range

Age group (years)	Number of Participants
25 - 44	2
45 - 64	7
65+	10

Table 2 Dysarthria Severity among Participants

Dysarthria Severity	Number of Participants
Mild	9
Moderate	4
Severe	4
Unknown	2

2.1.1 Participants

The criterion to include participants was that they should have dysarthria to the point where it affects their speech. We recruited 19 adults: ten males and nine females.

Their age ranged from 42 to 65+ years (Table 1). The severity of cases (Table 2) and participants' speech capabilities varied so that our system and our requirements could facilitate the entire spectrum of people with dysarthria. In the following section, we describe how we accounted for this disparity and range. We recruited all participants through charity organizations and social media.

2.1.2 procedure

Prior to undertaking the interviews, ethical clearance was obtained from users. Then, the participants completed a demographic questionnaire in which they had to indicate their age, location, severity of dysarthria, and preference regarding the method of communication or assistive device to be used in the interview.

We conducted the interviews during the first half of 2021. The timing is important because, during this year, much of the world was weathering the coronavirus, or COVID-19, pandemic. Since several countries, including many where we conducted interviews, were under lockdown for varying periods, we conducted all interviews online.

As all the interviews were conducted online and participant speech capabilities varied, each interviewee communicated their answers using different methods. Some relied on their own speech to communicate, whereas others, when unable to speak clearly or tired of speaking, either typed in the video call program's chat function or shared their screen to reveal their answers typed into a Microsoft Word document. Another group used an assistive device (e.g., a text-to-speech device) to communicate.

Several participants also had another person to assist them in communicating. Considering that the person helping them would be biased from their own experience, we wanted to clearly distinguish between the subjective feedback of the assistant and that of the participant. Consequently, we focused our questions and framed them to relate exclusively to the participants' experience.

Since the interviews were semi-structured, all participants were asked to answer a set of core questions, and then, depending on their responses, were asked follow-up questions. The first section was about their experience with dysarthria, including the history of their case, their speech style, and the associated effects on their lives at home. The second section focused on how they coped with dysarthria. They were asked about technologies or assistive devices they use and their experience using these. They were also asked whether they were using VHAs or any voice interface (VI) device/service. The last section was regarding the proposed system. In this section, the interviewer explained the proposed system concept and its functioning. Then, the users were asked about their feedback regarding the proposed system, voice cues that they would find convenient, and the system design they would prefer.

2.2 Data analysis

As the interviews were exploratory, we conducted a thematic analysis using an open coding approach, which is driven from the transcripts (bottom-up approach), utilizing the guidelines in [33]. The first author transcribed all interviews verbatim. We used NVivo 12 to assist with the coding process. We began by transcribing a few recorded interview files to generate the code logbook and then moved on to transcribing all files. We repeated the coding process several times to identify similar codes and refine these until we arrived at a final set of codes. Next, we validated the reliability of the coding process using inter-coder reliability to check the agreement between different coders regarding the data coded. Two of the authors coded randomly selected transcripts. Then, we evaluated the inter-coder reliability using Krippendorf's alpha, obtaining a value of.86. This method was selected as it is the most flexible one and produces a maximally accurate result [34], Moreover, researchers have been using it more often [35].

2.3 Results

In this section, we discuss the themes and the major points identified from the interviews with the study participants. The subheadings in this section are the themes we derived from our coding analysis. For each of these themes, we discuss the participants' responses and provide direct quotations from the interviews.

2.3.1 Tasks

Participants reported that they use VI in general and VHAs for different, specific tasks. These devices give the participant some independence. One participant commented that using voice assistants "would give my wife a break from me asking her to find that information for me." They mentioned various tasks, either those for which they were currently using the devices, or those for which they would like to use one if their health condition deteriorated. The majority reported that they most often used the VHA device to play music. Similarly, they noted that they used their VHAs to acquire information, such as like news, weather updates, and football scores. Furthermore, they also used VHAs for entertaining themselves-one participant states that they often ask their VHA to tell a joke, and another that they ask it to play the "Alexa, I love you" song. Interestingly, few participants indicated that they use the device for communication, such as to send messages and make calls. They also gave examples of tasks for which they were not currently using a device but would like to have the opportunity to do so. Examples are to control smart devices and home appliances, such as lights, heating, alarm systems, and curtains, and the volume of the television or music system.

2.3.2 VHA/VI experience

Participants shared their experiences regarding using VHAs or other voice interfaces, such as Siri or smart devices. Our first finding was that the ways in which each participant interacted with the devices varied. Some participants with mild dysarthria interacted with these directly through speech, whereas others with severe dysarthria used their cellphone to type the command they wanted to send to the device and then relied on their cellphone to speak it. This would occur with the help of either a text-to-speech application or an application relying on the user's stored voice, which is a recording made by the users when they were still able to utter phrases. These recordings are then transferred to synthesized speech. Another group used their augmentative and alternative communication devices to speak on their behalf. These devices are tools that aid communication for people with communication disorders. The device receives input through different input devices, such as a mouse, a keyboard, and a joystick, and converts it to speech.

The participants' experiences of interacting with the VHAs or other voice interfaces similarly varied, but the majority had negative experiences, regardless of the method they used to interact. When the participant interacted by speaking, the device often failed to understand their speech. One interviewee said, "I would rather type because nine times out of 10 you get the wrong answer." Their experience with the voice interface failing to recognize their speech drove them to use a different method of interaction that sent their input directly. Another participant also indicated facing the same problem when using the alarm feature, commenting, "I find myself yelling at them a lot. My alarm clock-I can't turn it off anymore." As dysarthria could co-occur with a physical disability, using a regular alarm clock or phone alarm requires the use of hands, which the user might be unable to do. In contrast, one participant observed that when the method of interaction was through a cellphone, the process was complicated, requiring more steps to send their command to the device. Moreover, another participant commented, "I suppose the only difficulty with... all these voice banking systems is that the pronunciation that you get from your synthetic voice is not necessarily the word that you want to come out, and therefore, Alexa does not always respond correctly. So, you have to learn how to type certain words in to get the correct pronunciation." This comment raises several issues regarding the effects of system design in VHAs and other voice interfaces on people with dysarthria.

Nevertheless, not every participant's experience was uniformly negative-some were happy with their interactions with their devices. These individuals most often interacted through their cellphones or through an augmentative and alternative communication device.

2.3.3 Proposed system design

Participants expressed a variety of perspectives regarding the potential of interacting with VHAs using non-verbal voice interaction. Overall, they responded positively to the idea. One group was interested in non-verbal voice interaction as a means to decrease their dependence on others for performing a task. Moreover, the technique would empower them to use other device features, especially if they only have limited use of their hands. Other participants compared it with other interaction methods, such as typing, believing that non-verbal voice interaction would be faster and more direct than typing. One participant noted that typing in order to interact with their VHA "feels backward."

Another group thought otherwise and disagreed that non-verbal voice interaction would be convenient. Some could not make any noise using their voice at all, and thus believed that the proposed system would not suit them. Moreover, the dysarthria of some participants was severe enough to make it difficult to distinguish between even non-verbal voice cues. Others noted that since they were physically able to perform all their tasks, they did not need to use their voices. One interviewee argued that using one's voice if one has dysarthria is fighting a losing battle. Another found typing to be faster, and because of the rapid deterioration of their condition, said that they preferred typing or using their eyes over using their voice.

The last group did not have a clear opinion. These participants had dysarthria that was still mild, and they believed that non-verbal voice interaction might be helpful were their condition to worsen

Sounds When asked about what sounds they could make-if they were able to make sounds at all-to be used as commands, participants were unable to give specific answers and found it difficult to decide on an exact sound. However, regardless of the severity level of dysarthria, most of them found vowels to be easier, albeit softer, sounds to make than consonants. A participant with severe dysarthria tried to enunciate the sound of a long form of the letter “E” (e.g., “eeee”), but their voice was breathy. The same participant found it challenging to enunciate combinations of vowels because their voice was similarly breathy when uttering all of them. Another vowel that was difficult for a participant with a mild case of dysarthria to enunciate was “U;” the participant had difficulty pulling their tongue back.

As for consonants, some participants were able to enunciate these, but these required more effort, and made the participants’ speech sound lazy or that the participant did not take the time to pronounce the consonant properly. Participants noted that the most difficult consonant sounds to make were “G,” “H,” “S,” “D,” “K,” “F,” “Ch,” “T,” “P,” “B,” and “L.” Two participants observed that words with multiple consonants, such as the word “consonant,” were more difficult to pronounce than words with only one consonant. For instance, a participant said that they found it easy to pronounce “T” and “L” separately, but that they found it much more difficult to pronounce the word “little.”

In summary, the ability of each individual to pronounce letters and words varies and is affected by various factors, such as the level, the cause, and the type of dysarthria. However, most of them found it easier to pronounce vowels.

List preference

When the participants were asked whether they preferred the system to have a predefined list of voice commands or whether they preferred to generate their own list, their opinions differed. Most respondents preferred to program their own voice commands. They argued that, since every case is different, voice commands should be personalized and customized. One participant indicated that creating their own commands helped make these easier to remember.

In contrast, a few participants preferred a predefined list. One individual stated that their choice was related to their age or generation, for they were used to “Plug and Play” devices. Another commented that developers would know the sounds that work best and, therefore, they preferred a predefined list. Another participant alluded to the same idea but said they would prefer to have the predefined list initially and then move on to create their own personalized commands. Many of the participants agreed that both options should be available to the user. For instance, one noted that they would use the predefined list to become comfortable with using the system and then generate their own voice commands, and another suggested that developers offer guidelines to start or provide a list of voices that the user can choose from.

2.4 Discussion

The objective of the interviews was to understand the users’ requirements; obtain feedback on the proposed interaction method, which involves non-verbal voice cues; and to determine the non-verbal voice cues that may be incorporated into the system. The interview results indicate that people with dysarthria face challenges in using VIs and VHAs and that these devices are not fully accessible. Difficulty is experienced both when interacting directly with the device and when interacting through an intermediary device. However, Ballati [2] indicated that people with mild dysarthria could use these devices to a certain degree. This finding is consistent with the results of other studies that address the issue of interactions between people with dysarthria and VHAs [36].

A person’s quality of life is a major issue, and we found that these devices improve different aspects of the quality of life. For instance, they give the user the ability to communicate. As per Light’s definition of communication [37], the act of communicating involves expressing needs, exchanging information, and engaging in social interaction. This was shown when the participants used the VHA to tell a joke, call a partner, or send text messages. This result supports that of prior research, which has indicated that a person with dysarthria could use music as a form for interaction with their children. Others with dysarthria use music to express themselves and to add a humorous effect when communicating with others [38]. Another aspect of improving the quality of life is that it gives these individuals independence in performing various daily tasks [39].

The participants in our study reported that they perform various tasks using VHAs. However, a highlight is that there is a limiting factor in the usage, in which the tasks VHAs were used for, is the ability of the participant to enunciate the correct commands. In other words, instead of searching for a task that would perform the specific job they needed, the users started to examine which tasks they were able to utter. For instance, if the user were able to enunciate “weather” and not “music,” then they would only use the VHA for “weather.” Regarding tasks, we also inferred from some users that there is a discoverability issue with VHAs. The participants sometimes did not know the capabilities of the device and how they could use it fully. This is an issue for all users and not specifically for people with dysarthria [40]. However, this would be more difficult for people with dysarthria, for they would find it challenging to discover the tasks verbally owing to their speech impairment.

In addition, we found from the interviews that people without physical disabilities could perform daily tasks themselves despite the severity level of their dysarthria. They were able to use their phones or other input devices if they wanted to use text-to-speech apps, perform a certain task, or find an answer to something. Moreover, if their case is mild and their speech is still intelligible, they will be able to use VHAs. However, their experience would not necessarily be similar to that of people with physical disabilities. Similarly, if the level of dysarthria becomes so severe that the participant cannot use their voice, they will obviously be unable to use VHAs. Consequently, the target audience of our system would be people with moderate dysarthria and a physical disability. We hypothesize that people who do not have a physical disability but have dysarthria will prefer to use voice rather than the alternatives that they currently use. We intend to test this hypothesis in another study. However, the audience is not limited to this group and could also include people with severe dysarthria as long as they are able to utter sounds using their voice.

To summarize, this study identified users' insights, experiences, and challenges with using VHAs. In addition, it identified participants' preferences regarding the list of tasks, sounds, and task-sound mapping. These data help to enhance the understanding of users and, thus, the ability to build more effective and more usable systems. The following section describes the design process of the system developed according to the users' needs and preferences.

3 Second Study: System Design and Evaluation

3.1 Design Elements

3.1.1 List

Creating the voice command list is a key part of the system design for these tools. In line with the study participants' responses about the list of commands they would prefer, we decided to offer users the customization option they preferred. However, the voice list itself will be controlled that is, we will provide users with a list of voice commands to choose from and then customize the mapping between the commands and the task to be performed.

This approach has been incorporated in the related literature. For instance, [23] collected preferred keywords–commands in their case-from participants and programmed these into the system. When the user starts using the system, they can choose various keywords that are easier for them, which are selected from the list provided by the participants in that study. In addition, [41] relied on the same process, allowing users to customize their list. Further, Hamidi [20] found that list customization resulted in improvement in accuracy rates, and the improvement was significantly greater for groups whose caregivers and therapists participated in customizing the list.

To study the impact of different list choices on users, we will compare this approach of partially customizing the list with that of providing a predefined list in the system. The lists in both approaches will be extracted from the interview results and users' preferences and capabilities, as in Parker's study [17]. In this research, the participants provided a list of appliances they wanted to be able to control using their voice. The researcher then selected the words pragmatically according to the functionality of the appliance (e.g., "on" and "volume"). After a participant recorded their voice, if the researcher found that certain voices were unclear, they replaced the unclear word with another word.

3.1.2 Sounds

Sound and mapping:

Mapping the sounds with their corresponding actions entails connecting each sound with a command. To have a usable, learnable, and memorable system, we created a framework for the sound-action mapping process (see Fig. 1). First, we set the criteria for selecting sounds. Next, we listed the sounds that met the criteria, decided on the mapping approach and, last, conducted the mapping.

In setting the criteria for selecting sounds, the initial criterion was users' preference. Almost all of the participants reported that vowels were convenient sounds for them to make. However, it was difficult for them to select specific preferred vowel sounds or any sound for the interactions. Therefore, we, as the researchers, had to make the selection. We adapted the selection criteria according to the findings of [42] and [20]. The first criterion was for the sound to be easy to utter, thereby lowering the likelihood of vocal fatigue. The second criterion was related to acoustic discriminability.

The first selection criterion, namely, the sound should be easy to utter, aims to lower the likelihood of vocal fatigue. Based on the users' preference, vowel sounds were selected, in addition to nasal sounds, which are easy for people with dysarthria to utter. We added nasal sounds to increase the number of command combinations. The second criterion is related to acoustic discriminability which is the ability to recognize different sounds [43]. The sounds of the vowels may overlap in some cases of dysarthria [44]. Therefore, we selected sounds that are in the corners of the IPA vowel chart to minimize and avoid overlapping sounds.

The second step was sound selection. Because of the disparate etiologies attributed to the different causes of dysarthria, various articulation capabilities emerged [45], which led to a limited number of vowels to choose from. For example, certain vowels (e.g., /i, a/) remain quite intelligible even in the individuals with severe dysarthria unlike other vowels [45]. Because of their capabilities, the vowel sound that we opted for comprised monophthongs, which are single-vowel voices. Notably, diphthongs (i.e., a combination of sounds) require changing the vocal tract configuration, which results in a steep second formant slope, a type of acoustic measure [46]. People with dysarthria find it challenging to pronouns diphthongs. Since a monophthong is composed of only one sound, it is less challenging to pronounce compared with diphthongs or other vowels.

The third and fourth steps, which overlap, entailed selecting the mapping approach and implementing it. As mentioned in the beginning of Section 3.1.2, the aim of the mapping process is to increase the usability and memorability of the voice commands by taking the users' preference into consideration. Researchers have applied several mapping approaches to map sounds with actions or controllers. For example, [30] and [42] used the tongue's position to map vowel sounds with the mouse movement and direction. Harada [47] incorporated a similar approach in a voice-driven drawing application. Norman [48] also discussed mapping as one of the design principles that aims to increase system usability. We followed the natural mapping design principle from Norman's design principles. Natural mapping takes place when the knowledge in our heads is integrated with the knowledge from the world around us. In other words, through natural mapping the relationship between our knowledge and what we are trying to control is clear and obvious.

We applied this principle by taking a concept in our daily life and applying it to our design. This could also be described using life metaphors. An example from our daily life is the iPhone brightness controller, which increases the brightness by simply sliding the control up. Another example is the volume button in phones or remote controls, wherein the upward direction represents an increase while the downward direction indicates a decrease. This metaphorical orientation (up means more, turn on, or increase) is not arbitrary; it results from physical and cultural experiences [49]. From this concept, we selected the /a/ vowel most of the participants were able to utter-, which is an open vowel that requires opening the mouth and positioning the tongue far from the roof of the mouth. In our mapping, this sound will represent open, up, raise, and increase. This vowel is mapped with commands that has increasing and turning o feature. So /a/ vowel is mapped with "Turning on Light" and "Increasing Volume." However, we added a nasal letter before the vowels (ma) to differentiate between the commands for increasing volume.

An effective design takes into consideration user behavior [48]. Accordingly, we used users' behavior as a basis for mapping one of our voices. We chose the commands according to the users' behavior or what they would say in certain situations. For the "Weather" command, we examined behaviors/spoken words that would be related to the weather. Then, we decided to use "oh" because it "is used to communicate the sense that something has 'just now' been noticed or realized" [50]. For example, a person could comment about the weather saying, "Oh, the weather is nice" or "Oh, the weather is cold." In her book [51], Diane described "oh" in the sentence "Oh, this weather is awful" as an attitudinal adverb that expresses emotion or attitude. From this, we used the /u/ vowel, which has the same sound, for the "Weather" command. Next, since people hum when they are trying to recall or repeat the lyrics of a song, for the "Music" command, we chose "Hmm." Table 3 summarizes the sound-action mapping list.

The last technique was extracting the voice from the words, following Haradah [47]. Who used the "ck" sound for the command "Click." In our mapping, we used the sound "ing" for the command "Ring," which is intended for calling someone. Similarly, the sound "am" was used for the "Alarm" command.

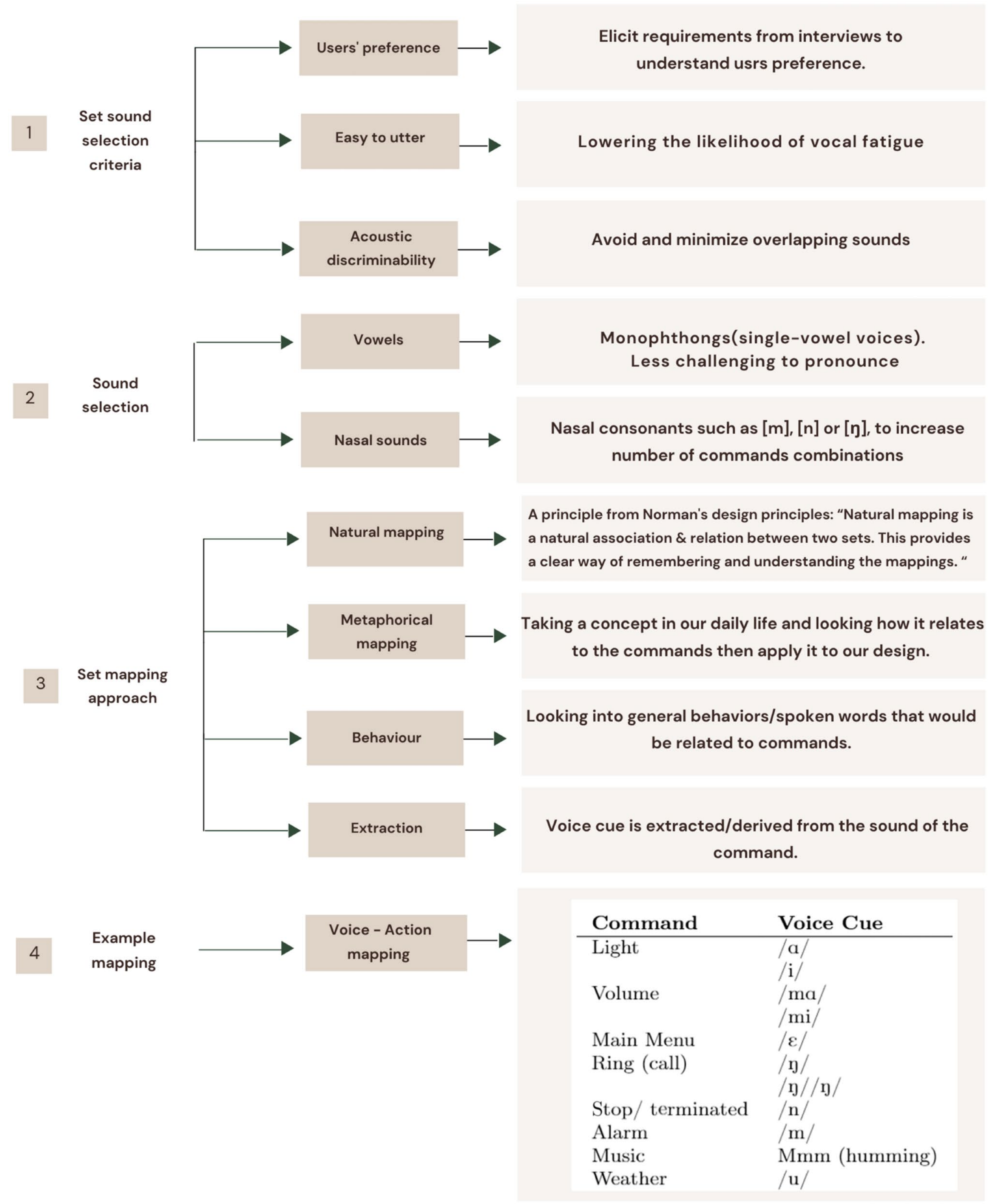
3.2 Initial Prototype and Preliminary Evaluation

In this section, we describe the prototype system based on our framework, and the preliminary testing we conducted.

The initial system comprised a list of five actions and five non-verbal voice cues. Each voice cue was mapped to one action, as shown in Table 4. The list was selected according to two factors: first, the interview results, which indicate the activities for which the participants use VHAs and, second, voice cues with only one sound, rather than, for example, commands with vowel and nasal sounds. The system was implemented on a Raspberry Pi that was connected to a microphone. Users must speak out a non-verbal voice cue to the microphone, and this command will be interpreted in

Table 3 Sound - Action Mapping

Command	Voice Cue
Light	/ɑ/
	/i/
Volume	/mɑ/ /
	/mi/
Main Menu	/ɛ/
Ring (call)	/ŋ/
	/ŋ//ŋ/
Stop/ terminated	/nÄ/
Alarm	/Äm/
Music	Mmm (humming)
Weather	/u/


1
Set sound selection criteria
Users' preference
Elicit requirements from interviews to understand usrs preference.
Easy to utter
Lowering the likelihood of vocal fatigue
Acoustic discriminability
Avoid and minimize overlapping sounds
2
Sound selection
Vowels
Monophthongs(single-vowel voices). Less challenging to pronounce
Nasal sounds
Nasal consonants such as [m], [n] or [ŋ], to increase number of commands combinations
3
Set mapping approach
Natural mapping
A principle from Norman's design principles: “Natural mapping is a natural association & relation between two sets. This provides a clear way of remembering and understanding the mappings. “
Metaphorical mapping
Taking a concept in our daily life and looking how it relates to the commands then apply it to our design.
Behaviour
Looking into general behaviors/spoken words that would be related to commands.
Extraction
Voice cue is extracted/derived from the sound of the command.
4
Example mapping
Voice - Action mapping

Command	Voice Cue
Light	/ɑ/
	/i/
Volume	/mɑ/
	/mi/
Main Menu	/ɛ/
Ring (call)	/ŋ/
	/ŋ//ŋ/
Stop/ terminated	/n/
Alarm	/m/
Music	Mmm (humming)
Weather	/u/

Fig. 1 Sound - Action Mapping Framework

the Raspberry Pi. The user's commands will be converted into text-based commands that will be sent to Google Assistant, which will process the request and produce a response (See Fig. 2).

At the end of the design and implementation phase, we conducted preliminary testing with seven participants. The aim of this user acceptance test was to verify the efficacy of the customized system that we built using the framework we proposed.

First, we tested the system with one participant to ensure that it performed well and had no bugs. In this test, we noticed some issues with the system, which we addressed before testing the system with the remaining six participants. Despite these issues, the participants gave positive feedback about the mapping and found it clear and easy.

3.3 Method

3.3.1 Participants

The main testing includes six participants who tested the system remotely, of whom three were male and three female, while three cases were mild and three were moderate.

3.3.2 Procedure

We conducted between-subjects testing, for which we divided participants into two groups. Group 1 tested the system using a set of pre-mapped commands-that is, we had already mapped the commands (voice cues) to the action desired. Group 2 tested the system using commands that they mapped to actions, and thus, they chose the command for each action.

The sessions started with five minutes of training, during which we explained the system to both groups and introduced the list of commands. To Group 1, we explained the process we followed to map the commands to the actions, along with the metaphorical concepts these represent. We asked the Group 2 users to map the commands to the actions, for which we gave them a set of non-verbal voice commands and a list of actions. Next, we asked participants from both groups to utter the commands at least once to ensure they were capable of doing so. After starting the system, we prompted the users to send it their commands. The command orders were listed in a random order, and the participants were instructed to follow any order that they preferred. Last, we conducted post-test semi-structured interviews to collect feedback from the participants about the system, specifically the sounds uttered and the mapping.

We asked the participants how easy the sounds were for them to utter and for their feedback regarding the system, particularly whether they would use it again and how simple they found it to use Their feedback and insights helped us to understand both system usability and user satisfaction. We also measured the effectiveness of the system by evaluating its success in performing the task and understanding the sounds.

Regarding the mapping, Group 1 was asked for their perspective about the mapping we provided. Since Group 2 had developed their own mapping we asked them to share the underlying reasons for this mapping. We also asked participants about their preferences on how they would like to use the system in the future whether it should be pre-mapped for immediate use or whether they would like to perform the mapping. The task success was evaluated as a sound/system effectiveness measure for both types of mapping. Moreover we emailed both groups about the mapping after 24 h to measure the memorability of the commands.

Table 4 Sound - Action Mapping

Command	Voice Cue
Light	/ɑ/
News	/i/
Ring (call)	/ŋ/
Music	Mmm (humming)
Weather	/u/

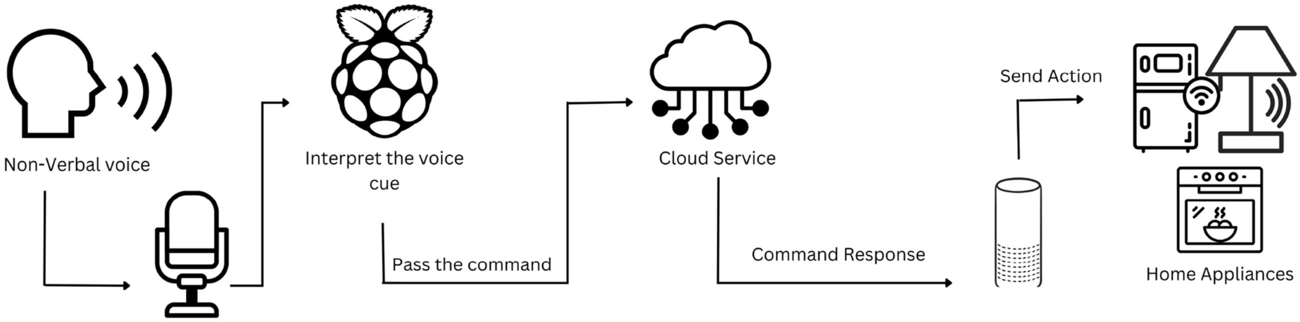


Fig. 2 System Architecture

3.4 Results

Through this preliminary testing, we aimed to test the efficacy of the proposed system. We found that the non-verbal voice cues selected were appropriately simple and utterable. All participants reported that they had no difficulty in uttering the sounds, that they found the system easy to use, and that they would like to use it again. One participant stated that the system was "easy to use because [it does not] use difficult letters like R, Z, or S."

We also found that vocal volume is important. Levels sometimes differ between voice cues, which is accentuated when the user does not increase their vocal volume for non-verbal voice cues. For example, two of the participants spoke more quietly when uttering "hmmm," and the system did not detect their almost voiceless commands. Similarly, one participant's voice was quieter when uttering "ing." Nevertheless, although they had to repeat themselves for the system to detect the commands, they did not find the sounds difficult to utter.

The results also show that Group 1 found the mapping provided by us to be learnable and memorable. This opinion is in parallel with the memorability measurements we used, in that all the participants remembered all the commands 24 h after using the system. When asked whether they would prefer to have the voice and actions pre-mapped, all the participants in Group 1 responded positively. When Group 2 were asked about the mapping 24 h later, only 40% could remember them, but all of them said they would prefer to do their own mapping. We also asked Group 2 about the reasons behind their mapping decisions. One participant indicated they had done the mapping randomly, whereas another stated they "tried to choose sounds that sounded a bit like the commands."

3.5 Discussion

The findings show that the system was successful and that the system that we built using the framework proposed has the potential to be used as a method of interaction, as an alternative to using different assistive technologies with varying modalities for people with dysarthria. The preliminary findings related to the question of mapping show that participants' opinions differ regarding mapping customization. Each group preferred to use the system the way they tested it, and this finding was unexpected. Group 1 found that the mapping made sense, whereas Group 2 preferred customization. A possible explanation for this variation is what Ellsberg [52] defined as ambiguity aversion. Participants chose the option they knew and thus avoided the risk of unknown factors. Therefore, further testing is required to determine whether user preferences for the pre-mapped and self-mapped systems differ.

4 Conclusion

In this study, we presented a method to design an alternative interaction mode with VHAs for people with dysarthria. We found that these individuals generally find it difficult to interact with VHAs and hence would find a faster, direct method of communication useful. The current data also highlighted that using non-verbal voice cues would be a convenient option. The framework we proposed in this study lays the groundwork for future research into non-verbal interaction using VHAs. This study thus contributes to existing knowledge on designing interaction systems by providing the steps we followed to design the sound-action approach.

Nevertheless, this study does have a significant limitation, namely, that only six participants were included in the preliminary testing process. However, this test was conducted just to verify the efficacy of the system. Hence, a larger sample should be included in a future test to obtain more quantitative and qualitative data and more insightful analysis on the specific utterances that should be mapped. Despite this limitation, the study certainly adds to current understanding about the research topic by providing information about the different experiences of the study participants, and their feedback.

5 Future Work

While the results of this study contribute to gaining insight into this field, further research with more participants is required to validate the results. Additionally, a broader range of evaluation measurements should be used.

Future work will also involve improvements to the system using a wider range of sounds. Sound combinations, rather than single sounds, should be introduced and measured in terms of how this affects system usability and the users' ability to utter them. A longer list of sounds would mean more actions could be performed, so the user would be able to use the system to perform more tasks. Additionally, a wake-up word could be added to the system to prevent it from detecting random sounds as commands. More customization options could also be added through which users could add the voices they prefer and map them to the actions they require.

Appendix A Interview Questions

1. About Dysarthria

(a) Can you tell me a brief history about your case with dysarthria?

(b) Could you describe your speech after having dysarthria?

(c) How do you feel when you are speaking?

(d) How does dysarthria affect your daily life?

2. Coping with dysarthria

(a) What technologies are you using to cope with dysarthria?

(b) Do you have/use a virtual home assistant? (If yes: What do you use these devices for? If No: Why you do not use it?)

(c) Would you list the top actions you use in these devices?

(d) How often do you use it?

(e) How did you find using it?

(f) What would you wish for virtual home assistants to have or look like?

3. The proposed system

(a) (after explaining the concept of non-verbal interaction) Do you think using non-verbal voice cues will be convenient?

(b) What non-verbal voice cues can be convenient for you to utter?

(c) Would you prefer having a predefined list of commands or having the ability to program your own commands?

(d) Are we willing to record non-verbal sounds to be used in building our system ?

Author Contributions All authors contributed to this research.

Funding This research did not receive any specific grant.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Abraham T. G., Kesavan R. K., and Raja S. K.: Feasibility of using automatic speech recognition with voices of deaf and hard-of-hearing individuals. In: Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, pp 373–374, (2017)
2. Fabio, B., Fulvio, C., Luigi De R.: Assessing virtual assistant capabilities with italian dysarthric speech. In: Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility, pp 93–101 (2018)
3. Enderby, P.: Disorders of communication: dysarthria. Handbook Clin. Neurol. **110**, 273–281 (2013)
4. Hartelius, L., Elmberg, M., Holm, R., Nikolaidis, s: Living with dysarthria: evaluation of a selfreport questionnaire. Folia phoniatrica et logopaedica **60**(1), 11–19 (2008)
5. Joy, N.M., Umesh, S.: Improving acoustic models in torgo dysarthric speech database. IEEE Transact. Neural Syst. Rehabil. Eng. **26**(3), 637–645 (2018)
6. Hux, K., Rankin-Erickson, J., Manasse, N., Lauritzen, E.: Accuracy of three speech recognition systems: case study of dysarthric speech. Augment. Altern. Commun. **16**(3), 186–196 (2000)
7. Luu, B., Hansberger, B., Chiu, M., Shivappa, VK., Karigar, GK.: Scalable smart home interface using occipitalis semg detection and classification. In: 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), pp 1002–1008. IEEE, (2018)
8. Mustafa, M.B., Rosdi, F., Salim, S.S., Mughal, M.U.: Exploring the influence of general and specific factors on the recognition accuracy of an asr system for dysarthric speaker. Exp. Syst. Appl. **42**(8), 3924–3932 (2015)
9. Yorkston, K.M., Strand, E.A., Kennedy, M.R.T.: Comprehensibility of dysarthric speech: implications for assessment and treatment planning. Am. J. Speech Lang. Pathol. **5**(1), 55–66 (1996)
10. Fried-Oken, M.: Voice recognition device as a computer interface for motor and speech impaired people. Arch. Phys. Med. Rehabil. **66**(10), 678–681 (1985)
11. Ferrier, L., Shane, H., Ballard, H., Carpenter, T., Benoit, A.: Dysarthric speakers' intelligibility and speech characteristics in relation to computer speech recognition. Augment. Altern. Commun. **11**(3), 165–175 (1995)
12. Thomas-Stonell, N., Kotler, A.-L., Leeper, H., Doyle, P.: Computerized speech recognition: influence of intelligibility and perceptual consistency on recognition accuracy. Augment. Altern. Commun. **14**(1), 51–56 (1998)
13. Rosen, K., Yampolsky, S.: Automatic speech recognition and a review of its functioning with dysarthric speech. Augment. Altern. Commun. **16**(1), 48–60 (2000)
14. Young, V., Mihailidis, A.: Difficulties in automatic speech recognition of dysarthric speakers and implications for speech-based applications used by the elderly: A literature review. Assist. Technol. **22**(2), 99–112 (2010)
15. Xiong, F., Barker, J., Christensen, H.: Phonetic analysis of dysarthric speech tempo and applications to robust personalised dysarthric speech recognition. In: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 5836–5840. IEEE, (2019)
16. Chen, F., Kostov, A.: Optimization of dysarthric speech recognition. In: Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.'Magnificent Milestones and Emerging Opportunities in Medical Engineering'(Cat. No. 97CH36136), **4**, pp 1436–1439. IEEE, (1997)

17. Parker, M., Cunningham, S., Enderby, P., Hawley, M., Green, P.: Automatic speech recognition and training for severely dysarthric users of assistive technology: the stardust project. Clin. Linguist. Phon. **20**(2–3), 149–156 (2006)
18. Hawley, M S., Enderby, P., Green, P., Cunningham, S., Palmer, R.: Development of a voice-input voice-output communication aid (vivoca) for people with severe dysarthria. In: International Conference on Computers for Handicapped Persons, pp 882–885. Springer, (2006)
19. Green, P., Carmichael, J., Hatzis, A., Enderby, P., Hawley, M., Parker, M.: Automatic speech recognition with sparse training data for dysarthric speakers. In: Eighth European conference on speech communication and technology, (2003)
20. Hamidi, F., Baljko, M., Livingston, N., Spalteholz, L.: Canspeak: a customizable speech interface for people with dysarthric speech. In: International Conference on Computers for Handicapped Persons, pp 605–612. Springer, (2010)
21. Derboven, J., Huyghe, J., De Grooff, D.: Designing voice interaction for people with physical and speech impairments. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational, pp 217–226, (2014)
22. An overview of the aladin project: Jort Gemmeke, Bart Ons, Netsanet Merawi Tessema, Janneke Van de Loo, Guy De Pauw, Walter Daelemans, Jonathan Huyghe, Jan Derboven, Lode Vuegen, Bert Van Den Broeck, et al. Self-taught assistive vocal interfaces. Proceedings Interspeech **2013**, pp 2038–2043 (2013)
23. Kim, S., Hwang, Y., Shin, D., Yang, C-Yl., Lee, S-Y., Kim, J., Kong, B., Chung, J., Cho, N., Kim, J-H, et al.: Vui development for korean people with dysarthria. J. Assist. Technol. (2013)
24. Ballati, Fabio, Corno, Fulvio, De Russis, Luigi: "hey siri, do you understand me?": Virtual assistants and dysarthria. In: Intelligent Environments 2018, pp 557–566. IOS Press, (2018)
25. De Russis, L., Corno, F.: On the impact of dysarthric speech on contemporary asr cloud platforms. J. Reliable Intell. Environ. **5**(3), 163–172 (2019)
26. Rudzicz, F., Namasivayam, A.K., Wolff, T.: The torgo database of acoustic and articulatory speech from speakers with dysarthria. Lang. Resour. Eval. **46**(4), 523–541 (2012)
27. Cai, S., Lillianfeld, L., Seaver, K., Green, J.R., Brenner, M.P., Nelson, Philip C., Sculley, D.: A voice-activated switch for persons with motor and speech impairments: isolated-vowel spotting using neural networks. Childhood **1**, 12 (2021)
28. Sporka, Adam J., Felzer, Torsten, Kurniawan, Sri H., Poláček, Ondřej, Haiduk, Paul, MacKenzie, I Scott: Chanti: Predictive text entry using non-verbal vocal input. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 2463–2472, (2011)
29. Igarashi, T., Hughes, J F.: Voice as sound: using non-verbal voice input for interactive control. In: Proceedings of the 14th annual ACM symposium on User interface software and technology, pp 155–156, (2001)
30. Harada, S., Landay, J.A., Malkin, J., Li, X., Bilmes, J.A.: The vocal joystick: evaluation of voice-based cursor control techniques for assistive technology. Disabil. Rehabil. Assist. Technol. **3**(1–2), 22–34 (2008)
31. Harada, S., Wobbrock, J O., Landay, J A.: Voice games: investigation into the use of non-speech voice input for making computer games more accessible. In: IFIP Conference on Human-Computer Interaction, pp 11–29. Springer, (2011)
32. Blair, J., Abdullah, S.: It didn't sound good with my cochlear implants: understanding the challenges of using smart assistants for deaf and hard of hearing users. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol. **4**(4), 1–27 (2020)
33. Braun, V., Clarke, V.: Successful qualitative research: a practical guide for beginners. sage, (2013)
34. Nili, A., Tate, M., Barros, A.: a critical analysis of inter-coder reliability methods in information systems research. (2017)
35. Guangchao Charles Feng: Intercoder reliability indices: disuse, misuse, and abuse. Qual. Quant. **48**(3), 1803–1815 (2014)
36. Fager, S.K., Fried-Oken, M., Jakobs, T., Beukelman, D.R.: New and emerging access technologies for adults with complex communication needs and severe motor impairments: State of the science. Augment. Altern. Commun. **35**(1), 13–25 (2019)
37. Light, J.: Toward a definition of communicative competence for individuals using augmentative and alternative communication systems. Augment. Altern. Commun. **5**(2), 137–144 (1989)
38. Kane, Shaun K., Morris, M R., Paradiso, A., Campbell, J.: at times avuncular and cantankerous, with the reflexes of a mongoose understanding self-expression through augmentative and alternative communication devices. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp 1166–1179, (2017)
39. Teixeira, António, Braga, Daniela, Coelho, Luís, Fonseca, J., Alvarelhão, Joaquim, Martín, Inácio, Queirós, Alexandra, Rocha, Nelson, Calado, António, Dias, Miguel: Speech as the basic interface for assistive technology. In: DSAI 2009-Proceedings of the 2th International Conference on Software Development for Enhancing Accessibility and Fighting Info-Exclusion, (2009)
40. Luger, E., Sellen, A.: Like having a really bad pa the gulf between user expectation and experience of conversational agents. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 5286–5297, (2016)
41. Hamidi, F., Baljko, M., Ecomomopoulos, C., Livingston, NJ., Spalteholz, LG.: Co-designing a speech interface for people with dysarthria. J. Assist. Technol. (2015)
42. Bilmes, J., Li, X., Malkin, J., Kilanski, K., Wright, R., Kirchhoff, K., Subramanya, A., Harada, S., Landay, J., Dowden, P., et al.: The vocal joystick: A voice-based human-computer interface for individuals with motor impairments. In: Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, pp 995–1002, (2005)
43. Wepman, J.M.: Auditory discrimination, speech, and reading. Element. School J. **60**(6), 325–333 (1960)
44. Lansford, Kaitlin L., Liss, Julie M.: Vowel acoustics in dysarthria: speech disorder diagnosis and classification. (2014)
45. Lee, J., Dickey, E., Simmons, Z.: Vowel-specific intelligibility and acoustic patterns in individuals with dysarthria secondary to amyotrophic lateral sclerosis. J. Speech Lang. Hear. Res. **62**(1), 34–59 (2019)
46. Kim, Y., Weismer, G., Kent, R.D., Duffy, J.R.: Statistical models of f2 slope in relation to severity of dysarthria. Folia Phoniatrica et Logopaedica **61**(6), 329–335 (2009)
47. Harada, S., Wobbrock, J O., Landay, J A.: Voicedraw: a hands-free voice-driven drawing application for people with motor impairments. In: Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility, pp 27–34, (2007)
48. Norman, D.: The Design of Everyday Things Revised and expanded, expanded Basic books, UK (2013)
49. Lakoff, G., Johnson, M.: Metaphors we live by. University of Chicago press, UK (2008)
50. Bolden, G.B.: Little words that matter: discourse markers so and oh and the doing of other-attentiveness in social interaction. J. Commun. **56**(4), 661–688 (2006)
51. Blakemore, D.: Relevance and linguistic meaning: The semantics and pragmatics of discourse markers, vol. 99. Cambridge university press, (2002)
52. Ellsberg, D.: Risk, ambiguity, and the savage axioms. The quarterly journal of economics, pp 643–669, (1961)