

T3: Requirement identification & Design Ideation

Project Title: HearMe

Team Name: HelpMeSpeak

Problem

People with speech impairments have a hard time communicating with the general public. For Dysarthria (a class of neurological speech motor disorder), in the beginning patients are able to communicate with people because people can understand them (though with difficulty). However, as the impairment progresses the communication becomes labored and eventually the patients reduce the communication attempts. As a result, the social network of the patients shrinks considerably. So much so, that only the caregivers of the patients are able to understand them for a very limited set of utterances. This reduces the self-reliance and autonomy of these patients. (Mackenzie, Paton, Kelly, Brady, & Muir, 2012)

However, this problem becomes exacerbated for subjects with motor impairments. Parkinson's disease, ALS, MS and brain injury patients, not only suffer from speech impairment but also motor control impairments. This makes the whole communication process really slow and prevents them from using any normal notion control based device which can enhance communication like computer. (Duffy, 2005)

Stakeholder Analysis

Primary: The end users are individuals with dysarthria. The people we interviewed are residents in a long term care facilities but it is conceivable that some potential users will live at home. Our target population has difficulty speaking due to a variety of health conditions, including stroke, Multiple Sclerosis and Parkinson's disease. (<http://www.asha.org/policy/TR2004-00262/>) They can and want to speak but it can be difficult for others to understand them. Due to their illnesses, many in our target group also have significant physical disabilities, and from our observations, they are confined to a wheelchair and possess limited motor skills. They have significant challenges with typical input/output devices, and are unable to grasp or move a mouse with control, cannot type with speed or accuracy, lack the strength and dexterity to use a touch screen effectively, and can have difficulty viewing a standard display due to visual impairment. They are eager to use technology to remain connected with friends, family, and caregivers, but often lack the right tools for their needs.

Secondary: Caregivers, family, friends and associates make up the secondary stakeholders. These people may have a direct role in helping the end users use the system, such as nurses or aides who helps the user to navigate to a room equipped with the system or help them connect or start up the system. From our ethnographic observations, we noted that the speech pathologist meets with the end users once a month or

so and there are times when this caregiver accomplishes the tasks of the end user if they are having trouble doing it. Family and friends can also perform this function, either in a care facility or at home. Secondary stakeholders can also be recipients of messages or interactions with the end users. This can include doctors, nurses, lawyers, insurance companies, friends and family. These interactions can be in-person, email messages, video conference events, or there could be documents created, sent and received by the user that involve these stakeholders. Some of these stakeholders may have roles in another category as a facilitator. A doctor could be a primary caregiver and an administrator, for example.

Tertiary: This category includes competitors, technology companies that may supply some of the components or software for the system and the companies affected by the success or failure of the software and other potential future users. If residents or patients at care facilities are able to successfully use a system like the one we propose, their quality of life will improve through increased connectivity with family and friends, a feeling of empowerment as they are able to do tasks on their own, and greater exercise of largely intact mental faculties trapped in unresponsive bodies. This can lead to better health for residents, higher satisfaction scores and create a better working environment.

The competitors who are dealing with the same target population could lose their potential customers if the software is a success. During our ethnographic observation, we found several competitors like Dynavox (www.dynavoxtech.com), Apple (www.apple.com) , Microsoft (www.microsoft.com) producing similar softwares for our target population. Technology companies that make key components could see benefits through increased sales from a successful system, such as specialty displays, sensing equipment or software. Since this software enables the users to communicate, the other enabling softwares too are affected directly by its success. Since the enabling software companies are dependent on this software, it could either be used by our target population or not used at all. If the software is a big success in the market, then in future, all the potential users may use this software for communication purposes.

Facilitating: These stakeholders include many members of the secondary group, as friends, family or caregivers will play a critical role in helping the end users with their use of the system. Also providing important support will be IT professionals who help get the system developed, installed and functional, and provide ongoing support for upgrades and maintenance. There is a session on group devices once every month in the home we visited where the technical staff customize the software for our target population.

Task Analysis

Check Email

Most users used email to communicate with family or friends. One user had a custom built system (by his family member). He communicated with the system by voice input and the system recorded a voice message (instead of text message) in the email. Other users were using Dragon Naturally Speaking or Windows Speech Recognition system to access, and write emails.

They used email on computers either in their own room or in the computer lab. People who had severe motor impairment used some form of speech recognition, while others used large button keyboards and on top rolling mouse for accessing the computers. They used the email from once or twice a day to a couple of times a week and they were previously trained on how to access the email based on their particular impairment.

The constraints for this task are bandwidth, power, computer access (in case the computers in the lab are all being used), physical and mental stamina because it takes time for our target population to send a single email.

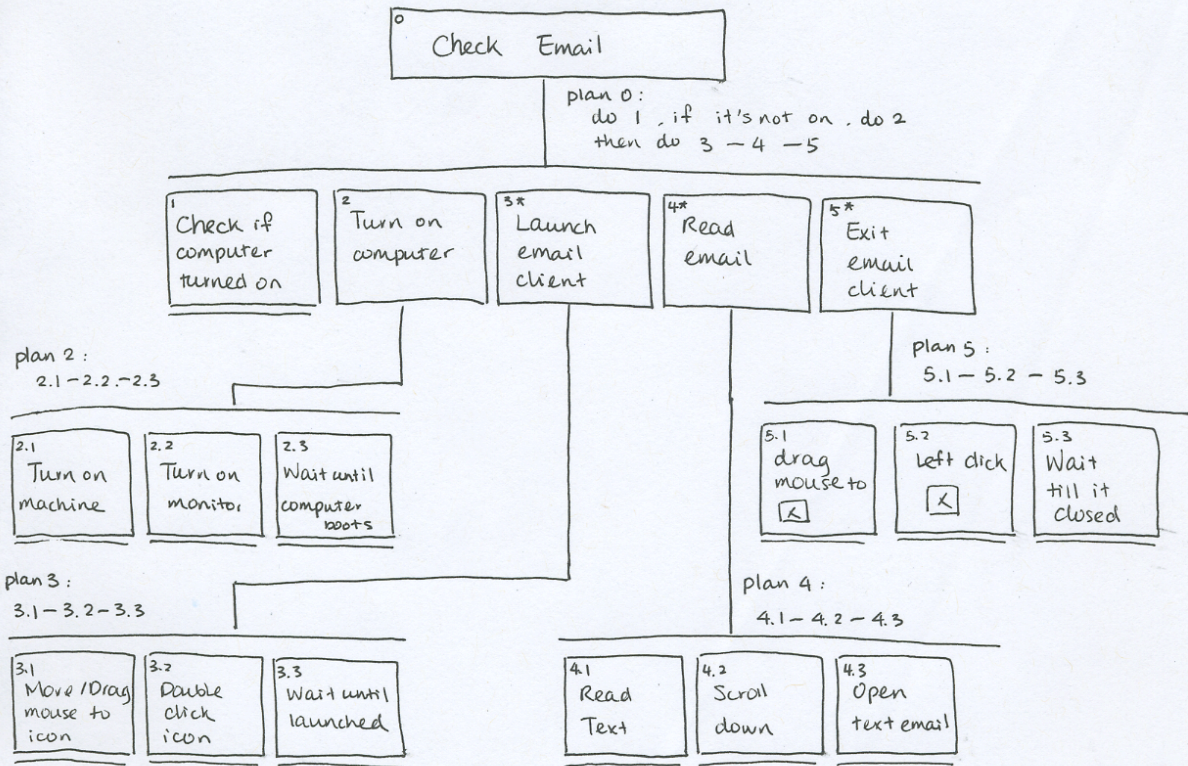
In case the computer was not turned on or they needed some other help (computer crashed, mouse fallen off the table, etc), the caregiver at the center helped them out, or if one of their family members were around they would help them. In one case we saw a social worker helping out the patient while the patient was giving a demonstration to our team member.

Pre-Condition:

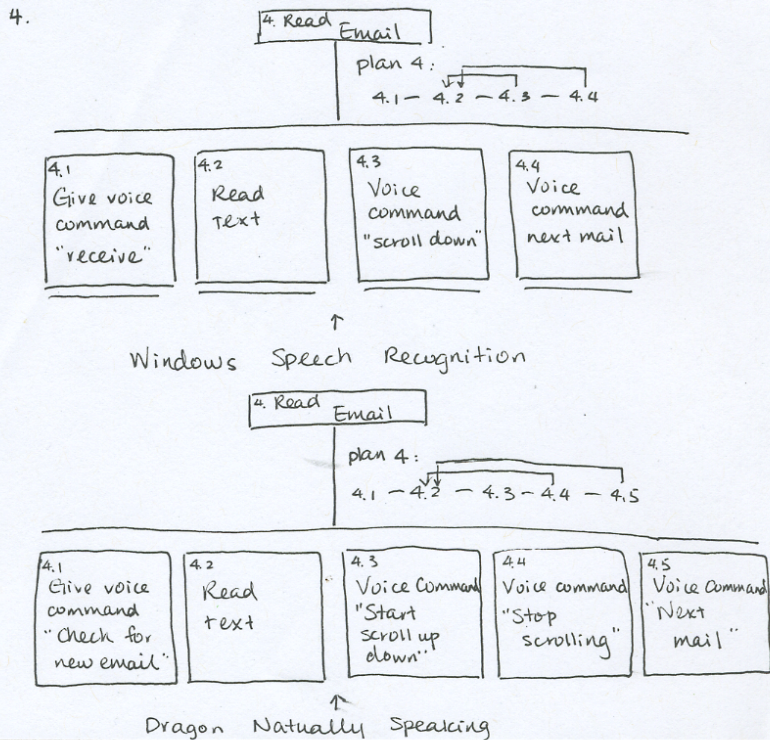
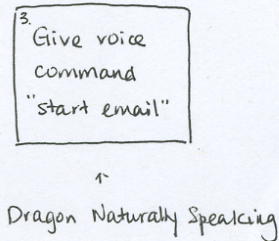
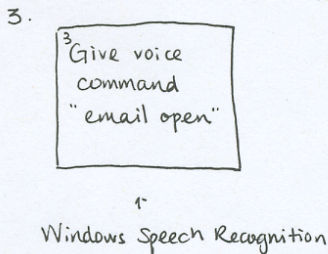
1. Have email account
2. Have computer access (in their room or in the computer lab)
3. Have internet access (for computer machines in the rooms the residents need to have their own internet connection, the facility does not provide internet connection)
4. Are able to see what they are interacting with (due to diminished eyesight the text and interface has to be large in size)

Exceptions:

1. Background noise effecting accuracy of speech recognition system.
2. Wrong input Syntax by the user.
3. System Crash
4. Extraneous input (while the user was conversing with someone else and forgot to put the speech recognition to sleep)



* step 3. 4. 5 have alternate ways to complete



5.

⁵
Give command
"close email"

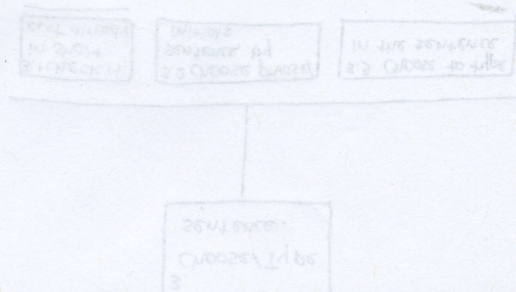


Alternate for Windows Speech Recognition

⁵
Give command
"click close"

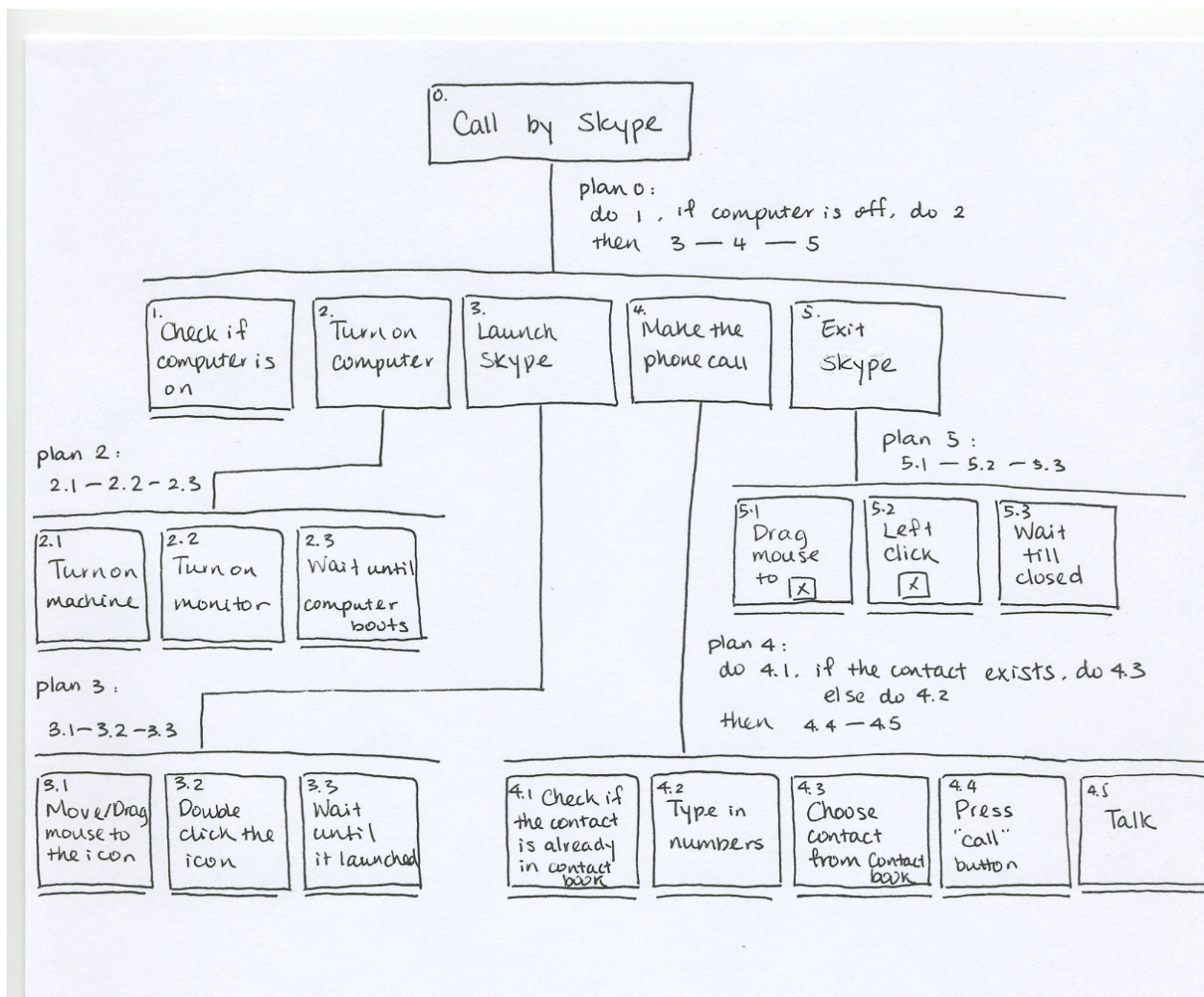


Alternate for Dragon Naturally Speaking



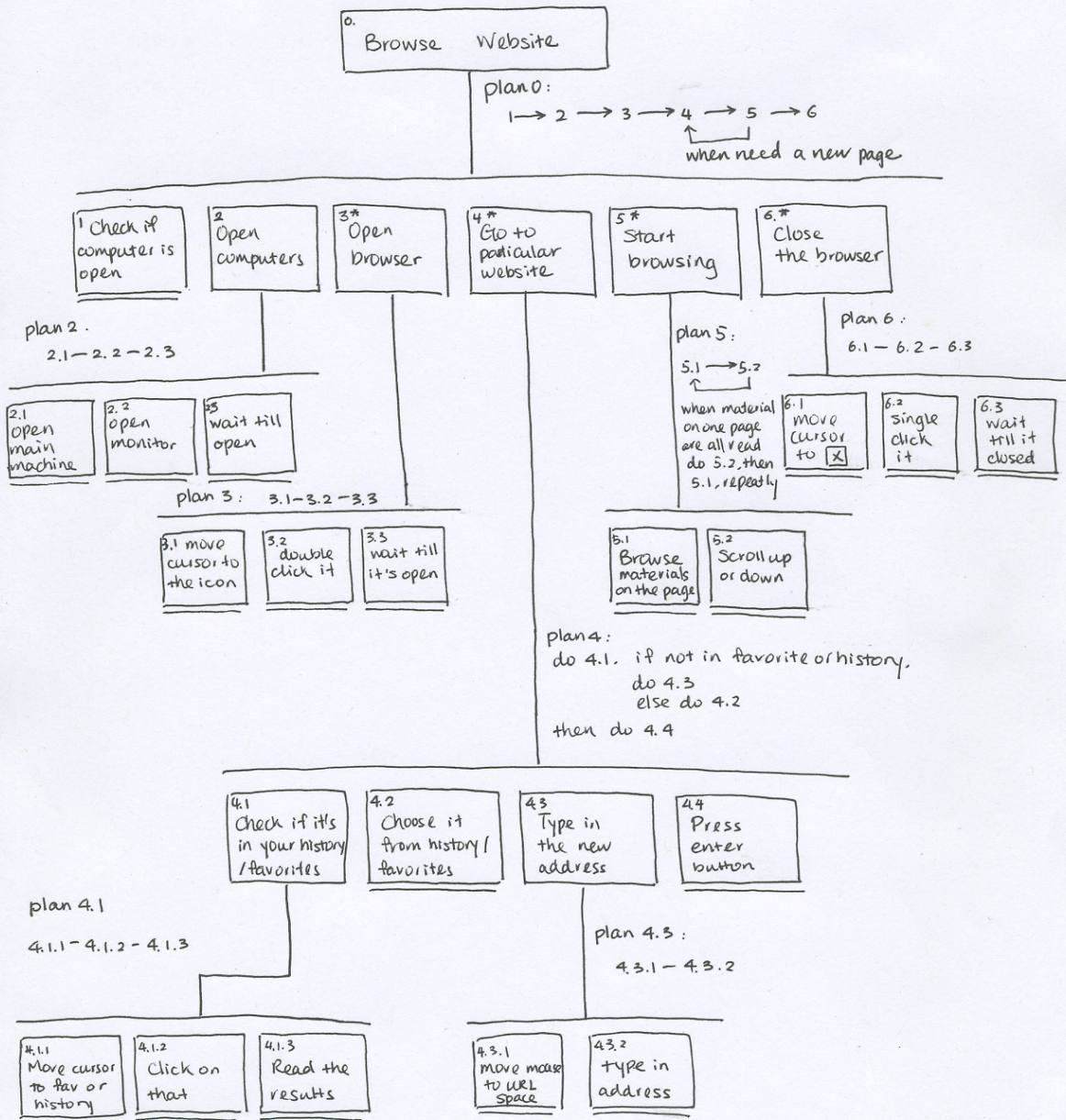
Use Skype

Quite a few of the users use Skype (www.skype.com) to talk with their family members and for this they use the computer lab or if they have a computer system set up in their room they use that. Three of our users whom we interviewed informed that they use skype to talk with their family, and also it was intimated to us by the speech pathologist during the initial tour of the facility. Similar to the email, skype was accessed using a speech recognition system or other input methods. They used the Skype once a week and they were previously trained on how to access the software based on their particular impairment. And this task has the same preconditions, exceptions and resource constraints as the previous task (email), with a slight change in the pre-conditions where now the user needs a skype account now instead of an email account.



Browse Websites

All of the users who access computers also browse websites and they consider this a way to get new information. This task (like Email and Skype) is either performed in their room on their own computers or in the computer lab. We got to know about this task while interviewing our target users and during the initial tour of the facility by the speech pathologist. The Speech pathologist later referred to it in other conversations as well. Caregivers and pathologist sometime also helps patients with some setups. Most of them knew how to browse website because they have been exposed to this technology before. Some residents there browse website everyday, while other do that at least one time a week. The pre-conditions, exceptions and resource constraints are again similar to Email. However, the only difference is in the pre-conditions where the user doesn't need a particular account, how he/she still needs his/her favorites set up in the internet browser software. This is because all the users we talked to have favorites set in their browser and they usually go to this subset of websites. For new websites, they get their family members or their caregivers (at the facility) to add a new website to the favorites.



* step 3. 4. 5. 6 have alternative ways to complete.
shown in another page.

3. alternate for person using "Dragon Naturally Speaking"

3
Give voice
command
"start internet
explorer"

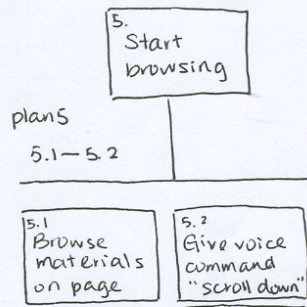
4. alternate for person using "Windows Speech Recognition"

4.
Give command
"go to address
....."
↑
speak address

alternate for person using "Dragon Naturally speaking"

4.
Give command
"go to (dictate
address here)"

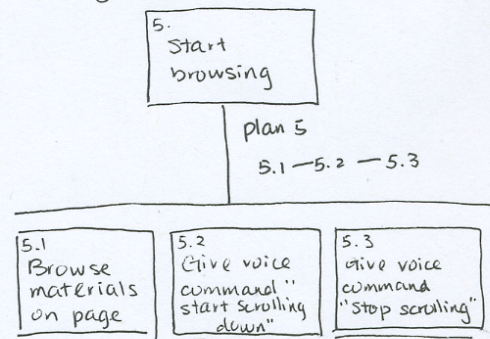
5. alternate for person
using "Windows speech
recognition"



6. alternate for person
using "Windows speech
recognition"

6.
Give command
"close internet
explorer"

5. alternate for person using "Dragon Naturally Speaking"

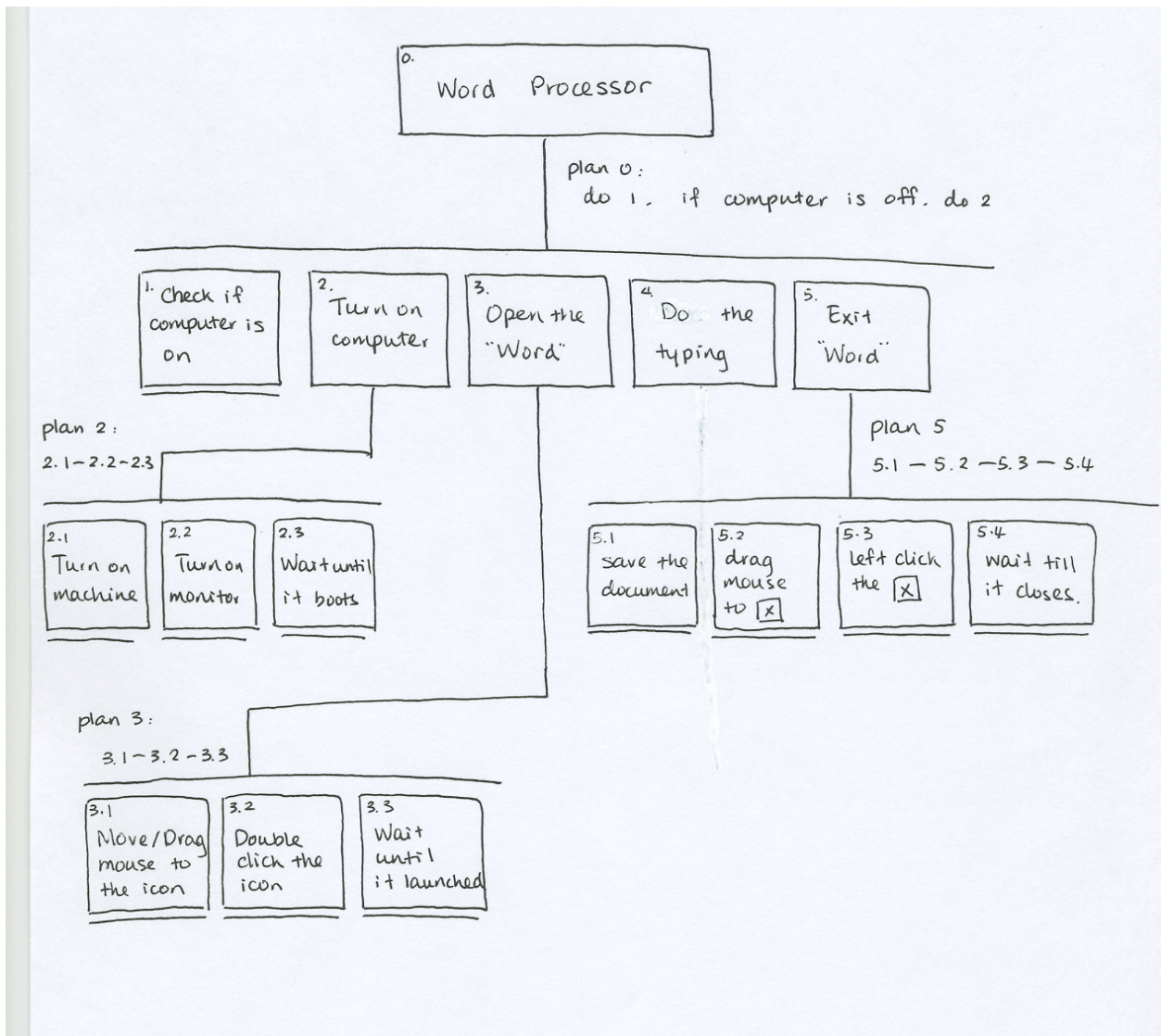


6. alternate for person using "Dragon Naturally Speaking"

6.
Give command
"click close"

Using Word Processor

Our target users use word processing software to write their diet charts and complaints to the facility. We were able to assess this in the interviews where the users gave us demo of how they used word processor. That is when we inquired the reason for using a word processor. To our knowledge this task was performed in the user's own room (we have no information if this task was performed in the computer lab). However, the precondition, potential exceptions and resource constraints are still the same from Email. However, the difference is that in the preconditions, they user does not require any sort of account online. Moreover, the caregivers and/or family members are additionally involved in getting the print from the printer, acquiring new printing paper and ink and loading the printer with these items.



Using Augmented Assistive Communication (AAC) Device

For our users who are unable to communicate with their voice or were very unintelligible, they use AAC devices to communicate with everyone around them. Our team members attended a Speech Device Group Session at The Boston Home (www.thebostonhome.org/) where they were able to observe some of the users of AAC devices at the facility. From our observation, using AAC device is an important part of their lives because they need to communicate with their caregivers and families and doctors, and lots of people rely on AAC device to do so. Some residents there also reflect that using AAC device to communicate has changed their lives. These users use their devices every day, all the time and everywhere; whenever they have to communicate. Every user is provided initial training on how to use the system and then once a month speech device group further adds on to the training of the users. However, the communication with the devices is usually very slow and requires a lot of patience on the part of the user and the listener (in this case caregivers, other patients at the facility and family members).

Preconditions:

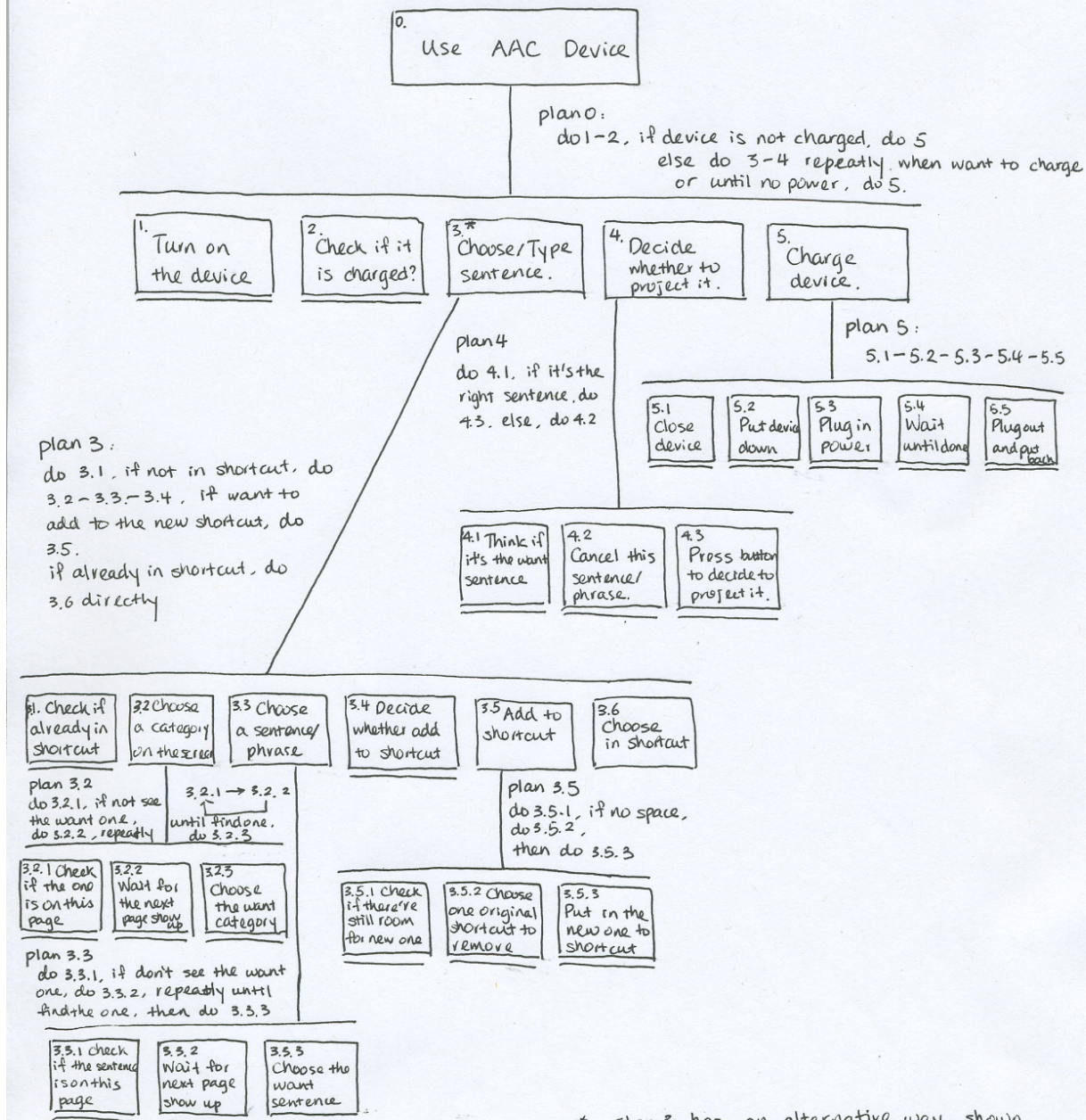
1. Training
2. Access to device
3. Device charged
4. Customization of the user interface for the user
5. Customization of the mounting on the chair and input modalities unique to the individual needs for the patient's disabilities.

Exceptions:

1. Inability to return to the top of the option tree.
2. Wrong input (for users with good motor skill: wrong word typed; for users with poor motor skill: control input button touched at the wrong time)

Resource Constraints:

1. Power / charging of the device
2. Availability of the device
3. Physical and mental stamina and the patience of the user and involved stakeholders.



*: Step 3 has an alternative way. shown in another page.

3.
Choose / Type
sentence

plan 3:

do 3.1. if already in shortcut. do 3.4
else do 3.2 or 3.3 depend on system you have

3.1 Check if
it's already
in shortcut

3.2 Choose
sentence
by initials

3.3 Type in
sentence

3.4 Choose it
from
shortcut

plan 3.2

do 3.2.1 - 3.2.2 - 3.2.3 - 3.2.4
until find one,
do 3.2.5

3.2.1
Choose
word or
phrase

3.2.2
Choose
an initial

3.2.3
Scan if it
is in the
current page

3.2.4
press the
next page

3.2.5
Choose
the phrase
/ word

plan 3.3

do 3.3.1 - 3.3.2
then if do 3.3.3 and have that in
prediction, do 3.3.4
else continue 3.3.2
3.3.2 → 3.3.3, then 3.3.5

3.3.1
Move hands
to type

3.3.2
Type in
words

3.3.3
See if your
ideal one
show in
prediction

3.3.4
Choose
from
prediction

3.3.5
Finish
type and
enter.

Using Assistive Phone to Make a Phone Call

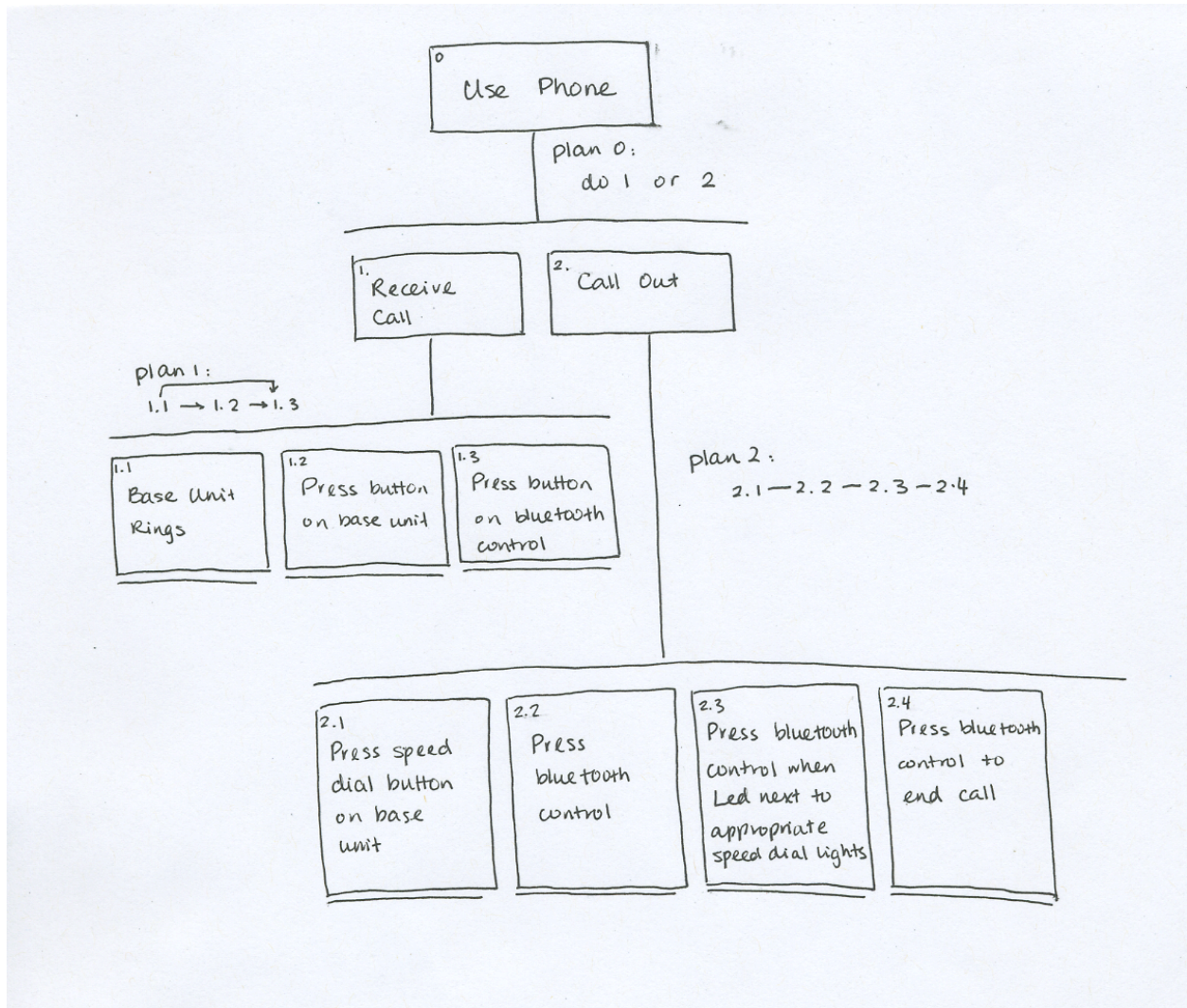
Each of the residents/patients at The Boston Home is provided with a special phone which allows them to call their family. The phone devices have large prints and straight lined buttons to use as speed dials (this phone does not have a regular keypad on the top). They phones come with a single touch bluetooth button which is attached to the patient's chair. When they have to call a family member, they just press the button on their wheel chair on time, and then the phone scan through all of the speed dial options one at a time. While scanning the LED of the current selected speed dial option is turned on. It starts with the first one, then the second one and so on. When the LED of the speed dial number, which the patient wants to call, turns on, the patient presses the control button on the wheelchair and the phone dials that number. Alternatively, for users who are not extremely disabled in the motor skills they just press the speed dial button they want to call. These phones are used couple of times a week in the room of the patient. They are provided training for this task and the only resource constraint is that the phone line is not cut. Besides the user the person who is being called in involved in this take.

Pre-Conditions:

1. Training
2. Phone installed in the room.
3. Speed Dial numbers programed.
4. Bluetooth control button installed on the wheel chair (for patients with severe motor impairment)

Exceptions:

1. Wrong input (for users with good motor skill: wrong button pressed; for users with poor motor skill: bluetooth control button on the chair pressed at the wrong time)



Functional Requirements

Functional Requirement #: 1

Description: The product should be able to recognize user's voice as main input, and consider the mouse control as auxiliary input.

Rationale: From our ethnographic observation, patients usually get frustrated when their voice commands are not recognized. Failure to comprehend the speech pattern will decrease user experience and prevent users from completing their tasks. The patient we interviewed usually kept repeating the command as many as three times and if it is still not recognized, he will give up trying it. If primary users' commands cannot be interpreted and there are no other auxiliary control methods, even secondary stakeholders such as caregivers and family members cannot help patients to complete their tasks. One patient we have interviewed, who is using the Gmail (www.google.com) voice system, can control the system by either voice command or the mouse. Even though he cannot use his hands to control the mouse, he told us his caregiver, whom we consider as the secondary stakeholder, could help him to complete the task by mouse

when his voice is not recognized by the system, and this has improved the efficiency. Thus besides the only input (voice) which can control the system to perform correctly, we should consider other control methods (mouse control) to make sure the system acts as the user wants even the voice command cannot be recognized. Since this will improve our primary stakeholders' interaction with electrical system such as computer, it will indirectly improve the communication between patients and secondary and tertiary stakeholders that are their families or caregivers.

Source: Patients with dysarthria and caregivers

Functional Requirement #: 2

Description: The product shall have both visual and audio output display.

Rationale: In our observation, the existing voice control system can misinterpret voice inputs quite often, especially from those people with speech impairment, whose voice might be different from normal voice. Misinterpreted commands or unintended commands may lead to incorrect execution or unwanted actions, such as changing or closing applications the user still wants to use. Meanwhile, we find that most patients with dysarthria have visual impairment, and it takes a lot of effort to see what is displayed on the screen even when using relatively large font sizes. And several interviewees said they made mistakes because they cannot see the small words or numbers in the computer system. Thus, both visual and audio displays are needed. This will benefit our primary stakeholders' interactions with the electrical device.

Source: Patients with dysarthria

Functional Requirement #: 3

Description: The system should allow the patients to control electronic devices without using hands.

Rationale: From our observation, patients with dysarthria are suffering from motor impairments, and usually they can only control one or two of their fingers; thus they have difficulties in controlling a mouse, typing in keyboard or controlling other devices such as iPad (www.apple.com) or Dynavox (www.dynavoxtech.com). Some patients are using specially designed mouse, which allows them to use a single finger to control the mouse, but it still need great effort to complete a task that needs multiple clicks or typing, because they need to adjust their wheelchair position to move their hands and arms. However, patients who don't have control of their fingers and can only use back of hands are still facing the problem of controlling a mouse or typing in keyboard. This greatly restricts their social communication options through computer or other electronic devices. With options like eye-gazing technology and voice control, end users can release lots of burden of moving their hands and arms around to control the mouse or type. One of our interviewers said, "The voice control Gmail system has significantly improved my life, because I can do things that other people can." This will also benefit secondary stakeholders because once our primary stakeholders having simple control of such as emails and Skype (<http://www.skype.com>), it will be easier to communicate to their families and caregivers.

Source: Patients with dysarthria and their family members

Functional Requirement #: 4

Description: The system should enable end users to know which mode the system is in.

Rationale: From our observation we found that patients with dysarthria have very weak voice and speak in very slow pace. We usually need to pay much attention to get what the patient is trying to say. And for patients who have been using the voice control system, sometimes they forget whether the system is still taking in their voice commands or has already stopped and started interpreting it. A common case is that when a user has finished saying his or her email contents and thinks the system has already stopped taking input, he or she may start talking to other people. However, if the system is still in “accepting mode”, it will end up with lots of unwanted materials in the email. Thus, letting the end users know what status the system is in and giving the user a signal of what the system is doing right now will make their commands to be more effective and avoid lots of unexpected mistakes. This benefits primary stakeholders most because it will improve the efficiency; however, it will also benefit the secondary and tertiary stakeholders since they will directly affected by the success of failure of the output of the system.

Source: Patients with dysarthria

Functional Requirement #: 5

Description: This product should be customizable for each end user.

Rationale: We find most patients with dysarthria are okay with short sentences but struggle with long sentences and while their voices are relatively strong at the beginning they become weaker and weaker as they speak a lot. Also, different users may have different accents, which make it difficult for database of a generic system to recognize different voices and their idiosyncrasies in tone and pronunciation. Meanwhile, the preferred command for a particular operation varies from person to person, for example different commands such as “strike xx” and “open xx” can be used by different users to open an application. One of our interviewees who was using a customized voice control system said, “Once I want to change my commands, I can just tell my brother to help me change it.” and his system appears to be running more efficiently than others because it is customized. Thus, having a customized database of commands for different users will benefit both the efficiency and correctness of the system. It will also make our primary stakeholders-patients-talk less and give better experiences using the system by having short decision trees which will makes the waiting time shorter. The process of customizing the system involves primary stakeholders that will record their commands; secondary stakeholders that will help the patients during the process; and the facilitating stakeholders that are usually the developers and technicians, which will introduce this system to both the patients and their families and caregivers.

Source: Patients with dysarthria and technicians

Functional Requirement #: 6

Description: This product should enable message transferring between end users and secondary users like caregivers, family members, friends and doctors.

Rationale: Due to their inability to say long sentences and their motor impairments, there are lots of situations where patients cannot express what they need when seeking help from caregivers or family members, as well as when they are having a consultation with a doctor or technicians. One interviewee told us sometimes she has difficulty in swallowing and cannot express that by herself, she will just drop her spoon on the floor to let her caregiver know this. She thought after several “trainings”, the caregiver would know that dropping a spoon means have difficulty in swallowing. However, this is not a good way to alert a caregiver to a potential dangerous situation. Having rapid interpretation of message and sending it to the secondary or tertiary stakeholders will greatly improve this situation. It will promote the communication between them and the patients.

Source: Patients with dysarthria, caregivers, family members, and technicians

Non-Functional Requirements

Non-Functional Requirement #: 1

Description: The product for dysarthria people should be platform independent and reliable.

Rationale: In our ethnographic observations, we noticed people using different types of system for communication. Some of the devices used by these people were PC, iPad (www.apple.com) or Dynavox (www.dynavoxtech.com). So our product would face constraints in deployment. Since all these systems have different OS and it is not backward compatible, we need a platform independent implementation to overcome this situation. Also we noticed that system crashes were highly rated as the undesirable by our target population. Thus the system has to be reliable for any tasks performed by the user. But we do not know the exact threshold level of our target users since it differs within our target population.

Non-Functional Requirement #: 2

Description: The proposed product's hardware requirements for visual display, keyboard layout, workstation layout, environment, display with reflections, display colors and non-keyboard input.

Rationale: People we observed were using specialized hardware for their devices since our target population is people affected by dysarthria, visual impairment, multiple sclerosis and Parkinson's disease (www.thebostonhome.org/). So hardware requirements form the backbone of our product and improve the accessibility. One person was using a specialized mouse for performing his actions on his PC. All the text on their devices had a larger text and color contrast was increased. This gives a good tip for designing our system. Ipad (<http://www.apple.com>) users wanted a stylus to operate their devices. Since the hardware costs are too high, we are not planning to alter or change their hardware. TV is used as monitor

for computer and people can switch modes between them.

Non-Functional Requirement #: 3

Description: The efficiency of our product can be measured in terms of response time and performance issue.

Rationale: We are dealing with speech-impaired people and they often try to communicate using their own voice. They wanted to use computers for basic tasks like web browsing and Skype (<http://www.skype.com>). Since user use his voice as input; response time for understanding their voice and completing the task forms the bigger part of the system. It should be able to clearly produce a result either a correct action or ask or provide suggestion to user. Since users are performing actions in a closed environment, other sounds are minimal compared to a public space. The system should not neglect the environment factors like noise and light. Since the target population is visually impaired, ambience should be properly set up before using the system.

Non-Functional Requirement #: 4

Description: The social organization and infrastructure of the care homes.

Rationale: This requirement deals with the enablers of the system. They are important for helping the target population to use the system and maintain the system. They record the user's difficulties with the system and provide feedback of the system. There are also in-house technical people for setting or customizing the product for these users. The users provide feedback to these people regarding the system and it can be any issue including battery replacing, charging issues, system standby problems, memory constraints. It happens once every month as observed by us, during the ethnography exercise. The social organization helps the users in completing their actions on the system. Sometime the caregivers actually perform the task of user on their behalf. This forms the first set of social circle the user interacts and they are trained to use system. Some systems are shared among users; they periodically use it according to their preferences. The end users have support services and communications regarding their system.

Non-Functional Requirement #: 5

Description: Button/Text size and color contrast in visual displays

Rationale: The system support the target population of speech as well as visually impaired people. So the concepts of color contrast and optimal visual display takes priority in developing the system. The visual displays like monitors or devices are either mounted on their wheelchairs at an appropriate ratio as observed during our field trip to the care home. The text size were very small in certain devices like Ipad(www.apple.com) since they were using in built keyboards. We need to incorporate larger text or visual displays into our project.

Non-Functional Requirement #: 6

Description: Circumstances in which the product will operate. Is internet necessary for the product to operate? Optimum light and power supply.

Rationale: The system is used by dysarthria or motor impaired people. So the system is usually mounted on a wheelchair or on wall. It varies with every single target user since it depends on the motor impairments. Some people use their mouth or head to operate the system. So we can expect the system to be controlled by user in any fashion. The system does not power-off as they cannot restart the system on their own. So we need the system to utilize the power efficiently. There is no dedicated internet access for the systems so we need to rely on the LAN access. There could also be rough handling of the systems and it should withstand the extra weight or heat exposed on the system. Since the home is air conditioned, we can expect optimum operating conditions for the system.

Design Concepts

Idea 1

System that is able to recognize and interpret the voice of a dysarthria speaker to initiate commands to a computer. Commonly available speech recognition software, such as that built into Microsoft's Windows (www.microsoft.com) operating system or available as free downloads, such as e-Speaking (www.e-speaking.com/), do not work effectively for people with dysarthria, and they have long navigation commands to start particular software and it is platform dependent. Every user needs to go through a list of options for completing a task and there is no shortcut available. It is neither compatible nor customizable and does not include proper interface to interact with other existing software in the device.

Our proposal would assist these users with speech recognition software. We are selecting our target population to use software to complete a task such as email or Skype (www.skype.com) or browser without using a mouse or keyboard. So our software is able to recognize the inputs from our target user group and it is platform independent. The user could save common instruction sets in the system and it would be able to recognize the persons' speech pattern. In situations where there is a common PC for many residents or patients to use, each user could store their voice profile to allow them easy access to the machine.

This proposed system would meet functional requirements in the following ways. Many of these people have other physical challenges and have difficulty with keyboards or using a mouse. They are frustrated by typing slowly or unable to grasp a mouse and thus it much easier for dysarthria speaker to use technology by voice enabled commands. The system produces legitimate visual or audio outputs of the commands being performed on the system and even reads out text if the user wants it. Since the user may want to continue with the command or can produce a different one. By producing these outputs we

can confirm the user action and could avert potential errors. There should be some very strong visual and audio clues to the user whether the action was successful or unsuccessful since some users may have visual impairments. The system provides clear information about the status of a user command being performed. We can use graphic icons to communicate to the user that it is still accepting the inputs or stopped. These graphic icons will help the user to understand the system in a better way and reduces the chances of making mistakes.

The system we propose also uses RFID tags attached to a person's wheelchair to facilitate login to our system. Since the system is often going to be put up in a public space, we need to authenticate the users and list down their favorite commands in a simpler way and RFID tag helps to achieve this. We allow each user to customize the system according to their needs and we prerecord the commands that he uses to accomplish a task. So each user can have different set of commands to invoke a program on the system. Since our system acts like an interpreter we can make the system to interact with other existing softwares. For e.g. the user can use our system to open Skype (www.skype.com) and start a conversation with his family or friends. So we customize the software for every end user according to his goals. The compatibility and customizing features with the existing softwares in the device are handled by our system and made simpler for the end user to invoke a task on it.

The software we are proposing will also meet those non-functional requirements. It is a simple speech interpreter and we are planning for platform independent one so that every end user is able to use the new software in his or her current device. The system should be compatible with the equipment available in the residence facility and be easy to install. The hardware requirements for the software remain unchanged, as we are not changing the user's device. For users who have a feeble or low audible speech might require an amplifier to project the voice into the system. The system should be reliable and robust in terms of performance. The user should not have to repeat their commands several times and it affects the efficiency of the system. This system could also be put in a public space since each user has his own voice profile. But we need to monitor the noise and amount of time the user takes to complete a task. The outputs are displayed in large text along with audio output.

An interaction metaphor is that our proposal could be viewed as a translator between a dysarthric speaker and standard voice recognition software or a computer. The standard systems are not able to process the altered speech patterns of our user population, much like a person who speaks English only cannot understand Russian or Chinese. To facilitate communication with a Chinese individual, an interpreter who speaks both languages can be employed to translate between parties. Much in the same fashion, our system would interpret the dysarthric speech from the user and make it usable by the computer.

Potential consequences include 1) misinterpretation, 2) inability to comprehend a speech pattern

and 3) assigning a pattern to the wrong individual. The system could misinterpret the speech of a user and perform the wrong function. If the user has to repeat them or alter the way they speak to get the system to respond as desired, the chances of adoption will likely decrease. If the system cannot recognize a speech pattern at all, the user may never be able to access the system. There is also the possibility that the system will assign the wrong profile to a speech pattern and present a user with the favorites and settings of another user. It may be difficult for a user to “tell” the system that they have the wrong profile. The system could also startup when the user is merely in close proximity to the unit and does not want to use the system. This error could present as an annoyance if visual and audio cues go off repeatedly. If multiple users are in the room (if, for example, the system is installed on a common computer), there could be conflicts on which profile to open or difficulty with interpreting commands if multiple users are speaking.

Idea 2

This system will move the mouse on the user screen using off-the-shelf (<http://www.tobii.com/en/eye-tracking-research/global/products/>, <http://www.asleyetracking.com/Site/Products/EYETRACPC/tabid/287/Default.aspx>) eye tracking technology. It will augment voice input for mouse clicks, turning on/off the mouse movement etc. The user will be able to zoom into (/out of) an area (box with the mouse pointer being the center) so that the text and buttons are further enlarged for the user. It will provide large visual feedback on the top of the screen to inform the user about the eye tracking status (on/off). Also the system will provide audio feedback (which can be turned on/off) for the options the user selects. Instead of using traditional voice recognition that according to our experience does not work well for people with speech impediments, this system will use pattern recognition that will be trained by the user at the start of usage. Inference details from ethnographic observation states the following options of using a computer: Microsoft Speech Recognition System (www.microsoft.com), Dragon Naturally Speaking (www.nuance.com) and custom built speech recognition system. All of these options use speech recognition, which to our experience is broken for people with speech impediments.

It takes the voice as an augmented input for the users to control some aspects of the system (click, turn on/off the mouse movement). All the while it uses eye tracking to move the mouse around the screen, which takes care of the need to avoid any motor input from the users. The feedback on the top of the screen about the eye tracking enabled (or disabled) will provide easy information about the current mode of the system. This acts as an audio/visual output from the system. This product could be customized for each end user based on their visual impairments and tastes. The click action could be customized with help of sounds like puff and so on. So each user can use the system based on their strengths. It also allows

the users to communicate with their family and friends without using hands because they will be able to type by using an on screen keyboard.

The hardware requirements for completing this system would require webcam and sensors. The environment setup for this system is simple and could be deployed in the existing devices used by our target population since these webcam are easily integrated and could be used with speech-impaired people. There are performance related factors that affects this system. If the end user is not able to concentrate on the screen for a longer time due to his physical strength then the system may produce undesirable results. The technical people form the social organization of this system. Since they train the users for using this system and periodically maintain the software and the related hardware. The system does not require Internet connectivity to perform the tasks and it is works on any device the user is currently using. The output produced by the system has large text to support the visually impaired users of our target group.

This system acts like a game console device for the movement of a player. Similarly the end user uses his eye to move up, down, left, right. Parts of the options are predefined and could not be changed or customized for a person while the 'click' or 'stop mouse' action can be customized. The game console helps the gamer to progress in a game and helps him achieve the goal. So does our system help the motor impaired people in helping them achieve their tasks through eye and voice based inputs.

Unintended sequences includes 1) if the user specifically does not turn on/off the eye tracking option the system will keep recording the movements even when the user turns around to talk to someone. 2) The user might give the voice instruction of "Click" but at that moment their gaze may move to a different option. This will result in a wrong button click. 3) And if the environment does not have enough light, then the eye gaze technology could result in errors.

Idea3

Virtual agent for the end users for talking to them and answering their questions. During our ethnographic observation, we found that the speech pathologist time is very limited and he could not spend enough time with a patient. He often was repeating the same things to every new patient coming to the home. So we can use a virtual agent to solve their issues and help them remember the features of the system. It can also be used to make a note of the sessions these users need to attend and provides periodic updates about the new features of the system. It can be called as a 'virtual buddy' to the end user.

This concept is relatively new to the field of health technology but has been shown to be effective in applications such as a virtual discharge nurse (Bickmore, et. al, 2009, Pfiefer & Bickmore, 2010). It is customized based on the end user and acts on their preferences. Recording the speech sessions and

helping them to remember things can further extend this concept. Another current concept like ‘Siri’ (copyright Apple, inc. www.apple.com) works well for normal people but could not be extended for speech impaired people. And it cannot be customized for the end user needs. It can just answer the question from web rather than picking the answers from pre-recorded sessions. A proper complete sentence is required for softwares like ‘Siri’ (copyright Apple, inc. www.apple.com) to produce a valid output.

This proposed system would meet functional requirement as the user communicates to the system by voice. Many of these people have motion impairments thus have difficulty with keyboards or a mouse. People who could create small or even pronounce single letter words can use this system. The system provides visual as well as audio outputs of the results if the user wants to. These results could be a search on the existing database or from web. And it can also provide additional feedbacks or his peers experience for the same problem. Since our target users usually have visual impairments, there should be some clues helping them whether their voice is been recognized and answered. We can use graphic icons to communicate to the user that it is still accepting the inputs or stopped. These graphic icons will help the user to understand the system in a better way and reduces the chances of making mistakes. The system will produce the results only after the user confirms the question. The interaction between the speech pathologist and end user can be termed and customized in the same way as the end user using this system. The systems answer all the user questions from the data provided by the doctors or caregiver or web. This message transferring could be like a dialog-based system. And also this system could be used without using hands since it is mainly controlled using audio input.

As to the non-functional part, the software we are proposing is simple virtual agent and could be deployed in any device. Users can just use their own device or the devices in a public place since the software is compatible with all the operating systems. The hardware is easy to install and don’t require any changes in users’ devices thus we do not require particular hardware support. As the same in Idea 1, for user’s weak voice, they might require an amplifier to project their voice into the system. The performance of the system should be stable thus user doesn’t need to repeat their commands. A simple PC or a device is enough for the project setup and the care givers can help the system by providing inputs about the task. The system will be operated upon user inputs and can produce legitimate text or audio as outputs.

The best interaction metaphor related to this system could be that of a company’s virtual customer care. Here we can search or ask information related to a product. It searches the company database and returns with a solution. It can serve many users at the same time and need not wait for a person to respond from the other end. The database is constantly updated and it pulls out the latest

information from the site. So our virtual agent in the same way helps the end users in training to use their device and also for proper usage of the device. The virtual agent while communicating an answer to the user could refer tips and user manual documents.

Potential unintended consequences are 1) the user's voices might not be recognized; 2) searching results isn't user's desired one because of so many options in the database; 3) system can't differentiate similar questions.

References:

Bickmore, T. W., Pfeifer, L. M., Byron, D., Forsythe, S., Henault, L. E., Jack, B. W., ... & Paasche-Orlow, M. K. (2010). Usability of conversational agents by patients with inadequate health literacy: Evidence from two clinical trials. *Journal of health communication*, 15(S2), 197-210.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management*. St. Louis, MO: Elsevier-Mosby.

Dix, A., Finlay, J., & Abowd, G. D & BEALE, R.(2004). *Human-computer interaction*.

Mackenzie C, Paton G, Kelly S, Brady M, Muir M., 2012, The living with dysarthria group: implementation and feasibility of a group intervention for people with dysarthria following stroke and family members, *Int J Lang Commun Disord*. 2012 Nov-Dec;47(6):709-24. doi: 10.1111/j

Pfeifer, L. M., & Bickmore, T. (2011, March). Longitudinal Remote Follow-Up by Intelligent Conversational Agents for Post-Hospitalization Care. In *2011 AAAI Spring Symposium Series*.

Storyboard

Idea 1:



Figure 1: Design Concept # 1

Starting from top left picture in Figure 1, the user enters the room and drives close to the computer. The computer is in stand by mode at this time. (Starting Bottom Left) User gives the voice command to the system to turn on and start an email. The system provides audio and visual feedback again all actions it takes. At the end of this interaction the user has sent an email to his family.

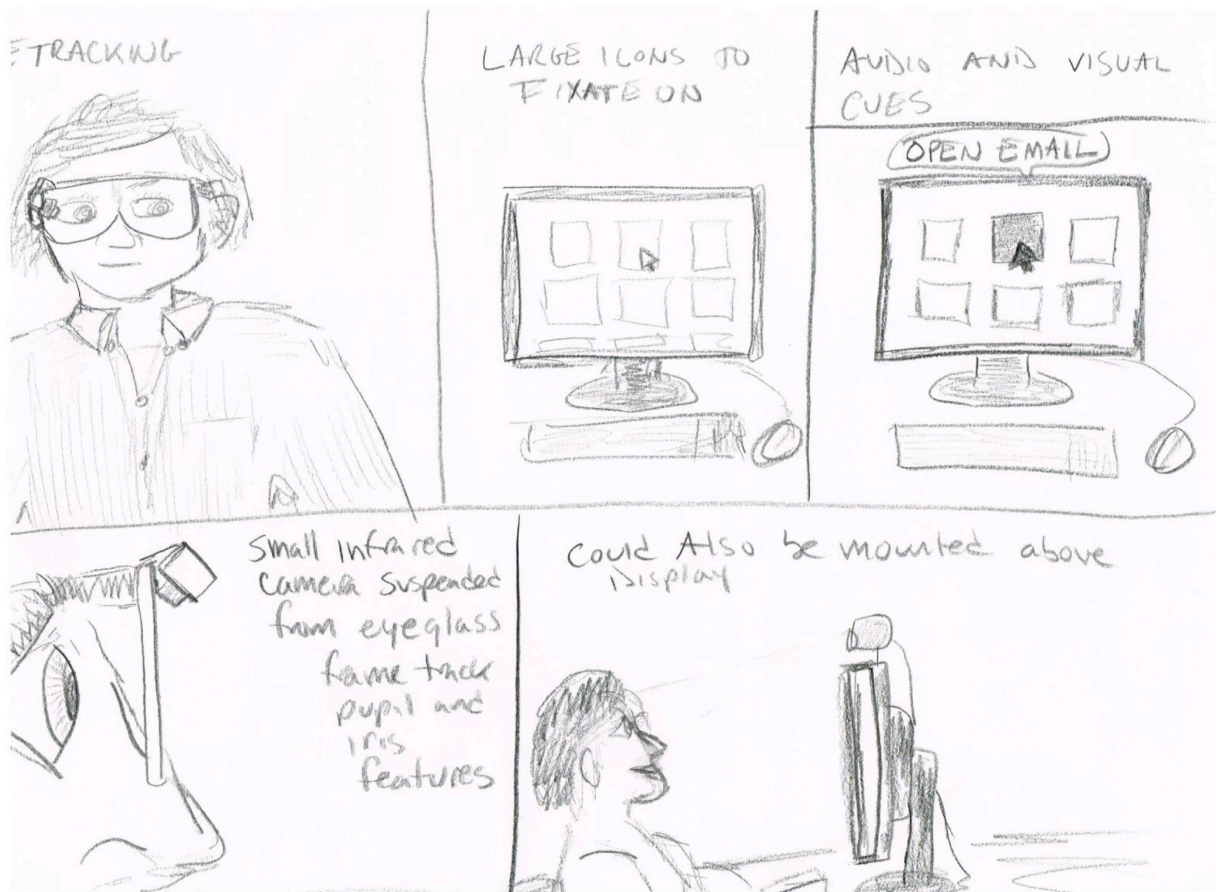


Figure 2: Design Concept # 2

Starting from top left (in Figure 2) the user is wearing an eye tracking goggle. Using this artifact the user is moving the mouse on the screen, which comprises of Large Icons to fixate on, and provides audio and visual cues to the user. Starting from bottom left, the eye tracker is described as a technology (infra red camera) mounted in front of the eye to track the gaze of the user. This technology can also be mounted on the display.

Idea 3



Figure 3: Design Concept # 3

This system will be a virtual agent that provides the user with speech impediments to perform regular speech exercises that he/she practices with the speech pathologists. In figure 3, starting from top left, the user Joan gets the virtual agent, and asks some daily life questions. Then they begin the exercises from where they left the last time. When the user is tired, she informs the agent, and the agent ends the session at this point. At the end Joan and Alex (the agent) bade farewell and the interaction ends.

Team Contribution

Problem Statement: Ganesh and Zhichun
Stake Holder Analysis: Steve
Task Analysis (minus task decomposition): Mansoor
Task Decomposition: Zhichun, Mansoor
Functional Requirements: Zhichun
Non-functional Requirements: Ganesh
Document Integration: Ganesh and Zhichun
Final Review: Steve

Individually we came up with 7 project descriptions (Ganesh 3, Mansoor 3, Steve 1, Zhichun 1) then short listed to 3.

