CÁTIA
AZEVEDO

Afluentia: SUPORTE À COMUNICAÇÃO PARA AFASIA FLUENTE

Afluentia: COMMUNICATION SUPPORT FOR FLUENT APHASIA

CÁTIA AZEVEDO

Afluentia: SUPORTE À COMUNICAÇÃO PARA AFASIA FLUENTE

Afluentia: COMMUNICATION SUPPORT FOR FLUENT APHASIA

"*The greatest challenge to any thinker is stating the problem in a way that will allow a solution*"

— Bertrand Russell

Universidade de Aveiro
2022

CÁTIA
AZEVEDO

Afluentia: SUPORTE À COMUNICAÇÃO PARA AFASIA FLUENTE

Afluentia: COMMUNICATION SUPPORT FOR FLUENT APHASIA

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica do Doutor Samuel Silva, Professor auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro, e da Doutora Ana Rita Valente, Investigadora do Departamento de Eletrónica, Telecomuicações e Informática da Universidade de Aveiro.

This work was supported by EU and national funds through the Portuguese Foundation for Science and Technology (FCT), in the context of the project AAL APH-ALARM (AAL/0006/2019)

o júri / the jury

presidente / president

Professor Doutor Paulo Miguel de Jesus Dias
Professor Auxiliar, Universidade de Aveiro

vogais / examiners committee

Professora Doutora Liliana da Silva Ferreira
Professora Catedrática Convidada, Faculdade de Engenharia da Universidade do Porto

Professor Doutor Samuel de Sousa Silva
Professor Auxiliar em Regime Laboral, Universidade de Aveiro

agradecimentos / acknowledgements

Esta dissertação vem encerrar o maior capítulo da minha vida até ao momento, o percurso escolar e acadêmico. Em primeiro lugar gostaria de agradecer ao meu orientador, Professor Doutor Samuel Silva, e à minha coorientadora, Doutora Ana Rita Valente, pela orientação, disponibilidade, paciência e incansável apoio e incentivo ao longo desde ano letivo. Agradeço também à Designer e às Terapeutas da Fala que participaram nas diferentes fases de avaliações, e sem as quais não teria sido possível alcançar os atuais resultados.

À Universidade de Aveiro e ao DETI por proporcionarem as melhores condições possíveis, ao IEETA e ao projeto APH-ALARM pela Bolsa de Investigação e a todos os professores, colegas e funcionários que se cruzaram comigo ao longo destes anos e que permitiram que desenvolvesse as minhas capacidades profissionais e pessoais.

Agradecer também a todos os meus amigos e à minha família, especialmente aos meus pais, por me terem incentivado a nunca desistir e me terem dado a possibilidade de estar aqui hoje. Por fim, e não menos importante, à Luna por me ter feito companhia e proporcionado os melhores momentos ao longo deste ano.

Palavras Chave

afasia, comunicação mediada por tecnologia, design centrado no usuário, comunicação aumentativa e alternativa, aplicação, terapeuta da fala, especialista em interação humano-computador

Resumo

A afasia é um distúrbio da linguagem provocada por danos cerebrais (por exemplo, acidente vascular cerebral) e que afeta a capacidade de comunicação de uma pessoa. Envolve diferentes graus de deficiência e pode-se manifestar por dificuldades em falar fluentemente ou dificuldade em encontrar palavras (anomia), mas também pode acarretar prejuízo na compreensão da linguagem falada, incapacidade de repetir palavras ou frases, deficiências na expressão escrita (agrafia), na compreensão da leitura (alexia) ou numa combinação de qualquer uma dessas dificuldades. Pode assim resultar em limitações na forma como a pessoa com afasia interage com outras para, por exemplo, exprimir como se está a sentir assim como as suas necessidades, impossibilitando que tenha uma vida mais independente ou tenha as suas dificuldades abordadas. Além disso, a afasia também tem um forte impacto na vida das pessoas ao redor do afásico (por exemplo, familiares, cuidadores), pois as dificuldades de comunicação da pessoa com afasia podem levar ao medo de deixá-la desacompanhada por aqueles que a rodeiam. Vários desafios surgem ao abordar as necessidades de comunicação de pessoas com afasia decorrentes da natureza diversa e idiossincrática da sua condição e embora ferramentas de comunicação assistida tenham sido propostas na literatura (por exemplo, usando pictogramas), as características da afasia geralmente tornam essas soluções parciais. Nesse sentido, este projeto foca-se na compreensão das características e necessidades dos pacientes afásicos e também na proposta de ferramentas de comunicação mediada por tecnologia que os abordem, no seu cotidiano. Este trabalho adota uma abordagem de design e desenvolvimento centrada no utilizador de modo explorar como as pessoas com afasia podem ser apoiadas na sua comunicação quotidiana recorrendo à mediação tecnológica. Foi assim conseguindo um design iterativo com desenvolvimento e avaliação de uma solução de prova de conceito para aspectos de comunicação, que foi progressivamente implementada e aperfeiçoada tendo em consideração os requisitos identificados e a avaliação contínua das soluções propostas, realizada com um grupo de foco composto por uma Terapeuta da Fala e um Especialista em Interacção Humano Computador. Após a obtenção de uma primeira versão do sistema, uma fase de avaliação com Terapeutas da Fala com forte experiência com pacientes com afasia também ocorreu de modo a entender e validar a aplicação alcançada, o que levou a mais fases de refinamento. Na sua atual fase de desenvolvimento, os resultados da avaliação mostram um bom nível de usabilidade e satisfação e definem o Afluentia como um terreno promissor para evoluir ainda mais a pesquisa em comunicação mediada por tecnologia de suporte a pessoas com afasia.

Keywords

aphasia, technology-mediated communication, user-centred design, augmentative and alternative communication, app, speech and language therapist, human computer interaction expert.

Abstract

Aphasia is a language disorder caused by brain damage (eg, stroke) that affects a person's ability to communicate. Involves different degrees of impairment and it can manifest by difficulties in speaking fluently or difficulty finding words (anomia), but can also entail impairment in spoken language comprehension, inability to repeat words or phrases, impairments in written expression (agraphia), in reading comprehension (alexia) or a combination of any of these difficulties. It can thus result in limitations in the way the person with aphasia interacts with others, for instance, to express how they are feeling as well as their needs, making it hard for them to have a more independent life or have their difficulties addressed. Additionally, this condition also has a strong impact in the life of those around them (e.g., family, carers) as the difficulties of communication, should anything happen, can lead to fear of leaving these patients unattended. Several challenges arise when addressing the communication needs of people with aphasia deriving from the diverse and idiosyncratic nature of their condition and although assisted communication tools have been proposed in the literature (e.g. using pictograms), the characteristics of aphasia often render them as partial solutions. In this sense, this project focuses on understanding the characteristics and needs of aphasic patients and also on the proposal of technology-mediated communication tools that address them in their daily lives. This work adopts a user-centered design and development approach to explore how people with aphasia can be supported in their day-to-day communication resorting to technology mediation. It was thus achieving an iterative design with the development and evaluation of a proof-of-concept solution for communication aspects, which was progressively implemented and refined having in consideration the identified requirements and the continuous evaluation of the proposed solutions, carried out with a focus group composed by a Speech and Language Therapist (SLT) and a Human Computer Interaction (HCI) Expert. After a first version of the system was achieved, an evaluation phase with Speech and Language Therapists with a strong experience with patients with aphasia took place in order to understand and validate the achieved application, which led to more refinement phases. At its current stage of development, evaluation results show a good level of usability and satisfaction and establish Afluentia as promising ground for further evolving the research on communication mediated by technology to support people with aphasia.

Contents

List of Figures

List of Tables

CHAPTER 1

Introduction

In this chapter, an introduction is made to the subject of this dissertation. In this sense, the contextualization of the theme and the respective motivation that led to this work's realization will be presented. It will also identify the main challenges and objectives that guide the study to be carried out. Lastly, a brief description of the document's organization and structure is explained.

1.1 Motivation

Communication can be described as the act or set of acts that allow a transmitter and a receiver to exchange information. Human beings have the great advantage of being able to communicate in many and diverse ways, whether through speaking, listening, writing, reading, transmitting sounds, signals, simply looking at each other or just the way of behaving.

Due to the great diversity of methods and forms of communication that human beings have, as well as the reach that they can have, it can be easily stated that communication is one of the greatest and most precious existing abilities. Seeing this ability in some way damaged or altered can cause abrupt changes on an individual's life [1].

Various events can wreak havoc on an individual's way of communicating. Brain injury is one of those events, which can give rise to several debilitating conditions, one of them being aphasia, and the most common cause of aphasia resulting from brain injury is stroke.

Aphasia is defined as an acquired language disorder resulting from brain damage. More specifically, aphasia is characterized by the existence of a multimodal reduction in the ability to interpret and formulate the meaning of language elements, manifesting itself in difficulties in understanding, reading and writing [2]. Although there are several ways to classify aphasia, many have as their first dividing factor the maintenance of fluency or not.

A communication impairment can severely reduce an individual's ability to socialize, maintain a job, attend school, or live independently [3]. Also, communication impairments can cause abrupt changes on an individual's personal and professional life, and also in the lives of those around them, whether they are family, friends or co-workers. That being said,

it is clear that giving people with aphasia an opportunity to manage appropriately all the communication changes due to their condition is urgent and a matter of mental health [1].

1.2 Challenges

It is believed that if general population knew more about aphasia, there would be greater understanding of the condition and of the communication needs when people with aphasia interact with others and consequent reduced stigma associated with the condition. Studies have shown that when asked about aphasia, most people only have heard of the word "aphasia", and some of them without really knowing what it means, often thinking that is something not even related to aphasia at all [4]. That being said, it is easily concluded that having heard the word "aphasia" not always means that people know what it consists of, what challenges people with that condition face and how to deal on a daily basis with aphasia diagnosed individuals.

Technology can be viewed as an answer for communicative challenges for people with aphasia. Nowadays, more than ever, technology has a huge role on people's everyday life, whether at work or at home and that is why the opportunity to link technology to the well-being of aphasia patients arises. Efforts have been made in this area in recent years, and although there has been an exponential growth of proposed solutions, few have become an effective or accessible solution for a large number of patients [5]. Most of them are computerized solutions and the mobile solutions that have been explored, despite the numerous advantages they have, in most cases, were designed regarding a set of conditions that affect communication, thus making these solutions comprehensive and not specific for individuals with aphasia [5].

When talking about assistive communication technology directed to people with aphasia, the term augmentative and alternative communication (AAC) devices comes up. AAC is used to describe all methods of communication other than verbal speech. Those methods vary and may be personalized to meet each individual's needs. Associating the use of technology with ACC is the challenge, as there is still a lot to be done and achieved, despite advances in recent years.

As recent studies show, computerized augmentative and alternative communication devices present a wide range of options, even though adoption to date has been limited, as industry leaders are estimated to only reach less than five per cent of individuals who could benefit from it [6]. That is due to the fact that such devices, as computerized AAC, present some disadvantages as for example being expensive because they require specialized hardware and software, and consequently are out of reach to many people [7]. Also, the fact that most of them have been designed with other populations in mind, such as individuals with amyotrophic lateral sclerosis, traumatic brain injury and multiple sclerosis [5]. Those who have been developed specifically for people with aphasia, even though in small number, are directed for individuals with severe aphasia, only being utilized when therapy to regain natural speech has failed [5].

With mobile technology increase and its many advantages, such has reduced cost, increased usability, network connectivity, embedded sensors, portability and population literacy with such devices [3], the goal is not only to associate AAC with technology but specifically with

mobile technology. In that matter, AAC mobile technology has been studied and designed in recent years, but, again, with other populations in mind, not reaching specifically aphasia diagnosed individuals in quantity. Many AAC solutions provide the possibility to transcript and synthesize speech as well as libraries of symbols associated to vocabulary organisations.

1.3 Objectives

The present dissertation aims to identify the communication challenges that people with aphasia face in their routine in order to characterize the motivations and specificities of the patient with aphasia and also to propose solutions that can assist these patients in their day-to-day communication needs. In that sense, there are certain goals that have to be met:

- Gather knowledge regarding aphasia and different existing assistive communication solutions.
- Adopt a User-Centred methodology, characterize the needs and motivations of aphasia patients regarding communication in everyday Scenarios.
- Define Requirements for each Scenario and model aphasia patients into Personas.
- Propose technology-based solutions for assisted/alternative communication according to the defined Scenarios.
- Achieve an iterative design, development and evaluation of a proof-of-concept solution for communication aspects.

1.4 Document Structure

The remaining of this document is organized as follows:

- **Chapter Background** aims to describe in a more precise way what are the challenges people with aphasia face due to lack of some of the most important communication skills in today's society, in what consists properly aphasia and what connection may have with augmentative and alternative communication techniques, devices and systems. A brief discussion is made and User-Centered Design is approached as well as some considerations regarding Supporting Development for Mobile Approaches are made.
- **Chapter Personas, Scenarios, Requirements** describes all different types of Personas, Requirements and Scenarios that were identified, properly justified. Some of methodologies used are also approached.
- **Chapter Iterative Development and Evaluation of a Low-Fidelity Mockup** describes the achievement, through an interative development, of a Low-fidelity prototype as well as defines some visual guidelines regarding the system.
- **Chapter Iterative Development of Afluentia** focuses on going into detail about everything that has been done and why, discussing the used methods to implement the defined requirements and exploring the system architecture as well as the technical options that were taken. An evaluation phase is described, carried out with Speech and Language Therapists (SLT) as well as with a Human Computer Interaction expert (HCI), which lead to the final version of Afluentia, also detailed in this chapter.

- **Chapter Conclusions** draws conclusions, as the name implies, from everything that has been studied and achieved and also describes what more needs to be done in this area related to future work.

Additionally, this document contains several appendices (Appendix A, Appendix B, Appendix C, Appendix D and Appendix E) to better understand some concepts, implementations and more detailed view of some information explored throughout the document. Appendix A contains the two types of standardized questionnaires, namely the System Usability Scale (SUS) and Post-Study System Usability Questionnaire (PSSUQ). Appendix B describes the explored technologies associated to the identified features and functionalities. Appendix C contains the first developed version of the user interfaces of Afluentia. Appendix D consists of an evaluation guide that was performed during the evaluation phase described at chapter 5 and Appendix E contains the final developed version of the user interfaces of Afluentia.

CHAPTER 2

Background

This chapter starts by highlighting the importance of communication in the human being and in today's society, and how aphasia affects it. After, the challenges people with aphasia face due to lack of some of the most important communication skills are explored as well as in what consists properly aphasia and what connection may have with augmentative and alternative communication techniques and devices, with support of relevant literature.

Finally, and given the challenges of developing for these audiences, a brief discussion is made and User-Centered Design is approached as well as some considerations regarding Supporting Development for Mobile Approaches are made.

2.1 Communication

Communication is a tool that allows any individual to express themselves, their thoughts, their ideas, their feelings and emotions. Oral language has always been considered by men the most distinctive feature of their superior nature in relation to other species, and that is why this function is given a state of superiority [8]. Thus, communication can be assumed as one of the most important skills human beings pursuit.

Communication can also be described as a process in which a transmitter and a receiver exchange information, which in the case of human beings is shared verbally or non-verbally by people [9]. Verbal communication is the use of words along with vocal noises in order to share information audibly, weather this process occurs in person or remotely (by audio call, video call, among others). Non-verbal communication can be best defined as a silent form of communicating with a person or interlocutor without using any form of speech. There are four important functions of non-verbal communication which are complement, regulate, substitute for, or accent a verbal message [9].

"During everyday communication, especially face-to-face interaction, vocal and visible behaviors are typically coordinated in ways that provide for their mutual performance. When people talk, they also locate their bodies, assume various postures, direct their eyes, perhaps

move their hands, altogether behaving in ways that constitute an interactive event" [10]. (Stanley E. Jones and Curtis D. LeBaron, 2002, p. 499)

When two interlocutors have their full communication skills intact, verbal and non-verbal communication complement, regulate and accent each other, as described by Jones and LeBaron. On the other hand, when verbal or non-verbal communication are compromised by any reason, one type of communication can substitute the other.

Having one form of communication altered, weather if its verbal or non-verbal, can compromise the way of living of an individual, even in the simple aspects of life. If verbal communication is altered in any form, asking for a coffee at a coffee house or asking where the toilet is in a grocery store can be compromised, being a teacher and having the need to transmit knowledge can be compromised as well as many other simple or not so simple tasks, sometimes severely reducing an individual's ability to socialize, maintain a job, attend school, or live independently [3].

Many conditions can cause harm to the ability of an individual to communicate verbally, one of them being aphasia.

2.2 Aphasia

"Aphasia is a disturbance of the comprehension and formulation of language caused by dysfunction in specific brain regions. It results from a breakdown of the two-way translation that establishes a correspondence between thoughts and language. Patients with aphasia can no longer accurately convert the sequences of nonverbal mental representations that constitute thought into the symbols and grammatical organization that constitute language" [11]. (Antonio R. Damasio, 1992, p. 531)

As stated by Antonio R. Damasio apahsia is caused by dysfunction in specific brain regions, that is, aphasia is a result of one or more episodes of brain injury. The disease processes that cause aphasia are acquired instead of congenital, that in most of the times is caused by stroke, which is estimated to be the most common cause of communication deficits in an older population [12]. There are several types of aphasia that manifest themselves in different ways and although there are also several ways to classify aphasia, two approaches have been more referred in the scientific community: the Luriana classification and the Neo-Associationist classification [2].

In Luria's approach there are two fundamental assumptions. The first results from the idea that lesions in different brain areas can cause changes in the same function. The second assumption is that an alteration is qualitatively different depending on the location of the lesion [2]. The Neo-Associationist approach assumes that linguistic behavior has a direct relationship with specific brain areas. Injuries to cortical structures can result in Fluent and Non-Fluent Aphasia and injuries to subcortical structures can result in Subcortical-type Aphasia [2]. In this dissertation the Neo-Associationist approach will be the used one since it continues to be very popular among clinicians.

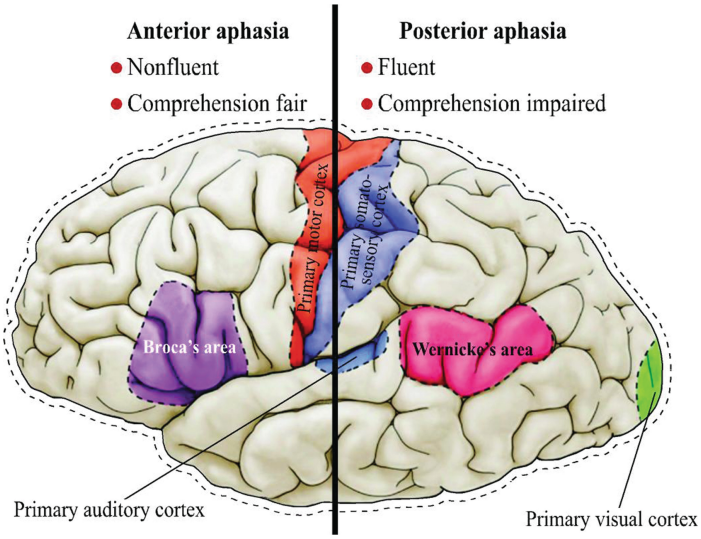


Figure 2.1: Anterior and Posterior Rolando's fissure brain, resulting in Non-Fluent and Fluent types of Aphasia. Image used from the book *Clinical Neurological Examination and Localization* by Vinit Suri [13].

As stated, in the Neo-Associationist approach, injuries to cortical structures can result in Fluent and Non-Fluent Aphasia. As Figure 2.1 depicts, Fluent Aphasia presents a brain lesion on the right hemisphere, located in the posterior region of Rolando's fissure and Non-Fluent Aphasia presents a brain lesion on the left hemisphere, located in the region anterior to Rolando's fissure [2].

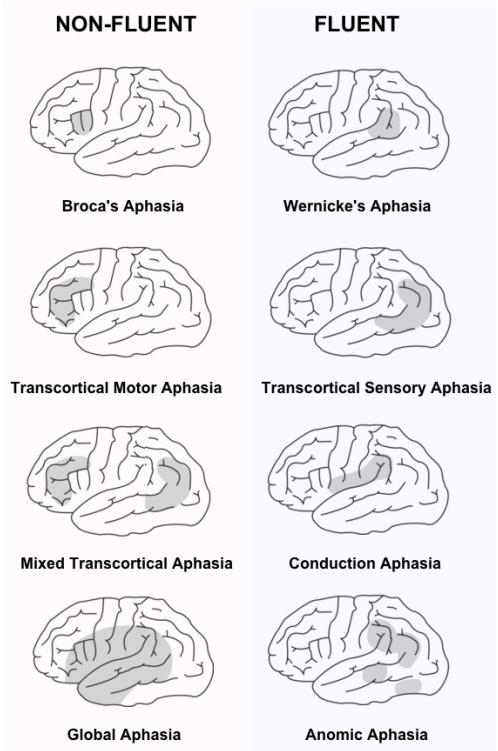


Figure 2.2: Damaged brain areas that result in the different types of Fluent and Non-Fluent Aphasia. Image used from the website *PsychDB (Psychiatry DataBase)* [14].

As illustrated in Figure 2.2, the different types of Fluent and Non-Fluent Aphasia are associated with a specific injured area of the brain, whether located in the anterior or posterior region of Roland's fissure.

In Fluent types of aphasia there is an ease in the production of long sentences combined with a marked difficulty in finding the correct words (anomia). Since there are no articulation problems, speech flows more easily, but the content of the message may lack meaning. Fluent Aphasia is divided into Conduction Aphasia, Anomia Aphasia, Wernicke´s Aphasia and Transcortical Sensory Aphasia. In Non-Fluent types of aphasia there are mostly production and articulation problems however, a listener can still understand what the speaker is trying to say. It is divided into Broca's Aphasia, Global Aphasia, Transcortical Motor Aphasia and Mixed Transcortical Aphasia [2].

All those eight types of Aphasia are identified taking into consideration seven aspects of communication that may be impaired, being them Fluency, Naming, Oral Language Comprehension, Simple Sentence Comprehension, Repetition, Reading and Writing [2].

	Conduction	**Anomia**	**Wernicke´s**	**Transcortical Sensory**
Fluency	Maintained	Maintained	Maintained	Maintained
Oral Language Comprehension	Maintained	Maintained	Altered	Altered
Repetition	Altered	Maintained	Altered	Maintained
Naming	Altered	Altered	Altered	Altered
Reading	Altered	Maintained	Altered	Altered
Writing	Altered	Maintained	Altered	Altered

Table 2.1: Different types of Fluent Aphasia and how the different language domains are affected by each of the types [2].

As it is possible to see in Table 2.1 and when speaking of Conduction Aphasia, there are four aspects of the ones mentioned above that may be altered, being them Repetition, Naming, Reading and Writing. On the other hand, when speaking of Anomia Aphasia only Naming may be altered with the possibility of minor changes on Reading or Writing. In Wernicke's Aphasia all aspects of communication are altered except for Fluency while in Transcortical Sensory Aphasia all are altered except for Fluency and Repetition. One common aspect of the four types of Aphasia mentioned is that all of them have their Fluency maintained (Fluent types of Aphasia).

	Broca's	Global	Transcortical Motor	Mixed Transcortical
Fluency	Altered	Altered	Altered	Altered
Simple Sentence Comprehension	Maintained	Altered	Maintained	Altered
Repetition	Altered	Altered	Maintained	Maintained
Naming	Altered	Altered	Altered	Altered
Reading	Altered	Altered	Maintained	Altered
Writitng	Altered	Altered	Maintained	Altered

Table 2.2: Different types of Non-Fluent Aphasia and how the different language domains are affected by each of the types [2].

Looking at Table 2.2 and regarding Broca's Aphasia, all aspects of communication might be altered except for Simple Sentence Comprehension while in Global Aphasia all aspects of communication are in some way damaged. Concerning Transcortical Motor Aphasia, Fluency and Naming are altered. In Mixed Transcortical Aphasia, only Repetition suffers no harm. One common aspect of the four types of Aphasia mentioned is that all of them have their Fluency altered (Non-Fluent types of Aphasia).

Be the bearer of one of those eight conditions gives people no other choice than therapy, but, given the importance of communication in every day routine, will it be enough to live independently while the process of recovery or rehabilitation is not accomplished? That is when the term Augmentative and Alternative Communication (AAC) joins the discussion.

2.3 Augmentative and Alternative Communication

Augmentative and Alternative Communication (AAC) is a set of techniques, methods and strategies that have the purpose of augment, complement, or replace speech of individuals with complex communication needs [15] [16]. AAC strategies are generally categorized as aided or unaided, with aided further divided into high- and low-tech and unaided also called no-tech [6].

Unaided or no-tech AAC techniques do not make use of any external equipment and involve different kinds of basic tools such as gesture, gaze, facial expression, body postures, sign language and any kind of voluntary motor movements in order to deliver non-verbal messages [6]. However these strategies can be used anywhere, do not break, are free and the fact that no-tech AAC is considered the oldest of the three categories given its reliance on the interpretation of facial expressions, their use is limited as stated by Michael Williams, an AAC user:

"Gestures can get you a cup of coffee in the morning, but they do a poor job of telling your friend about that delicious piece of cake you had the other night. Gestures can only express things in the here and now. [They] are poor candidates for expressing things like truth and beauty" [6].

Aided AAC, weather it is low-tech or high-tech, involves the use of external devices. Low-tech AAC supports are simple aided techniques that do not need any electronics or electricity and include books and display boards with extended lexicons of images and phrases to aid the communication process, cards, writing and drawing [6]. On the other hand, high-tech AAC devices involve computers or electronics. Devices falling under this category, such as smart devices, dedicated AAC devices, custom hardware and software devices and smartphones usually take the form of an electronic symbol-based dictionary or tend to rely on the translation of a user's intended meanings into speech via speech generating devices [6] [15].

Speech and language therapists typically try low-tech AAC solutions first since the use of simple display boards and communication books is both cost-effective and easy to get. Furthermore, the high costs and difficult training required by most high-tech AAC devices may limit access to high-tech AAC and hence the utility of speech generating equipment [15].

"In turn, an optimized use of high-tech AAC should be researched to provide a faster means of communication, in comparison to low-tech, by prioritizing the communicative needs of the users over the needs of the system" [15]. (Y. Elsahar, S. Hu, K. Bouazza-Marouf, D. Kerr, and A. Mansor, 2019, p. 2)

2.3.1 Augmentative and Alternative Communication for Aphasia

Individuals with aphasia are part of a group of people with complex communication needs, and since Augmentative and Alternative Communication is intended to increase, complement or replace the speech of those individuals, it can be part of the solution regarding recovery and well being of an aphasia person.

As stated before AAC strategies are generally categorized as aided or unaided, with aided further divided into high- and low-tech. In this chapter, and with the crescent use of technology by the common citizen and considering the main goal of this dissertation, high-Tech AAC solutions will be explored. Table 2.3 shows some of the main solutions developed in recent years regarding high-Tech AAC. None of the thirteen solutions presented were designed specifically for aphasia users but to every kind of user who may take advantage of AAC.

	Designed for Aphasia?	Year	Custom Hardware?	Synthesized Speech?	Price?
Dynamo	No	2001	Yes	No	1767€
Dynamyte	No	2000	Yes	Yes	5212€
Vantage	No	2001	Yes	Yes	?
I-Series	No	2015	Yes	Yes	11618€
Accent 1400	No	2020	Yes	Yes	6710€
Cboard	No	2018	No	Yes	———
Quick Talk AAC	No	2011	No	Yes	18,27€
iCommunicate	No	2010	No	Yes	49,99€
SymboTalk - AAC Talker	No	2018	No	Yes	0,99€/item
TouchChat HD - AAC	No	2019	No	Yes	149,99€
Sono Flex	No	2010	No	Yes	99,99€
Spoken - Tap to Talk AAC	No	2019	No	Yes	11,48€/ month
LetMeTalk	No	2013	No	Yes	2,49€/item

Table 2.3: Explored High-Tech AAC Solutions and the different characteristics identified and associated to each one of them.

The first three solutions presented are *Dynamo* and *Dynamyte* by Tobii Dynavox and *Vantage* by Prentke Romich Company and were innovative in terms of AAC technology as they were among the first to emerge in this area in the early 2000s. The three of them present a combination of custom hardware and software.

Dynamo is not a Text-to-Speech device as it uses a dynamic display and digitized speech, that is, speech which can be recorded directly into the device. It can develop sentences using single words or it can use combinations of phrases organized in themes to express ideas.

A more advanced *Dynamyte*, released after 2000, has almost all of the hardware and software capabilities of the *Dynamo* but has a dynamic screen, which can display a hierarchy of pages and communication is facilitated by a built in word prediction program and a dictionary, made accessible from any page. It uses synthesized speech, that is, computerized speech or text-to-speech features.

Vantage is the first all-touchscreen device from the Prentke Romich Company. Its main features are: Audio recording and speech output; Synthesized speech capabilities; Icon Prediction to simplify locating and learning the icon sequences; Spelling and Word Prediction functions; Digitized speech to add sounds, songs, or additional languages; Auditory prompts to aid individuals with visual impairments.

None of the previous three devices are currently commercialized, but both Tobii Dynavox

and Prentke Romich Company still exist and continue to work bringing High-Tech AAC solutions to the paradigm, most of them being a continuation of the first devices. As in addition to developing software for mobile technologies, such as the case of Tobii Dynavox that developed *Sono Flex*, they also continue to have their own devices with customized hardware and software such as for example *I-Series* in the case of Tobii Dynavox and *Accent 1400* in the case of Prentke Romich Company.

I-Series is an eye gaze-enabled speech generating device featuring the world's leading eye tracker. This Windows-based device is controlled completely with the eyes. *Accent 1400* offers a wide choice of vocabulary options to suit the diverse needs of augmented communicators and it is a text-to-speech device as it generates speech. Both of them are customized hardware devices.

Regarding applications for mobile devices, there are many options, one of them being *Cboard*, that is only available for Android and was released recently, in 2018. It aids communication with symbols and text-to-speech and with more than 3400 symbols from the Mulberry Symbol Set, it is possible for the user to create his own custom boards for different situations in life. *Cboard* also comes with support for 33 languages.

Quick Talk AAC app is based on categories of pairs of buttons to select phrases and words (100 maximum). 11,000 images and symbols from the Smarty Symbols library are available and it uses a recorded voice or text-to-speech functionality. It was designed to help individuals who are non-verbal to communicate through images and symbols. It is available for iOS and Android mobile devices.

iCommunicate is exclusive for iOS mobile devices and is a visual and text AAC application as well as task completion and audio visual prompting. Allows the user to create pictures, flashcards, storyboards, routines, visual schedules and record custom audio in any language and comes preloaded with 10,000 SymbolStix (symbols from the Smarty Symbols library) pictures.

SymboTalk - AAC Talker has predefined communication boards from different areas of life, each board containing symbols or images. Clicking on a symbol displays it aloud and adds it to a sentence which can also be displayed out loud. It is available for iOS and Android mobile devices.

TouchChat HD - AAC is a symbol and text-based AAC tool only available for iOS. It is a communication solution that allows non-verbal individuals to create words, phrases, and messages using synthesised or digitised speech. *TouchChat* can be customised to suit individual needs, has a library of over 10,000 SymbolStix, four pre-programmed vocabulary organisations, and uses spelling with basic word prediction. Text-to-speech voices are available.

Sono Flex is an AAC vocabulary app that turns symbols into speech. It offers language to non-verbal users who are not yet in full control of literacy. Provides text-to-speech voices and fully customized buttons, grids and folders and it is only available for iOS mobile devices.

Spoken - Tap to Talk AAC is a natural language assistant that utilizes state-of-the-art machine learning predictions as every word selected brings up the words the user may want next. The more it is used, the better it gets at predicting what the user wants to say and

utilizes next-generation text-to-speech. It is available for iOS and Android mobile devices.

LetMeTalk enables the user to line up images in a meaningful way in order do display that row of images as a sentence using Synthesised Speech. The image database of *LetMeTalk* contains more than 9,000 images from ARASAAC (AAC Symbols and shared resources) and the user can add existing images from the device or take new photos with the built-in camera. Supports several languages and allows for switching back and forth between languages. It is available for iOS and Android mobile devices.

2.4 Discussion

After the analysis performed in the previous subsection, 2.3.1, there are some considerations that must be made and that are relevant to this dissertation.

First of all, all the solutions were designed taking into consideration non-verbal individuals, weather they face Autism, Down Syndrome, Amyotrophic Lateral Sclerosis, Aphasia or even Apraxia, providing them an AAC tool. Even though non-verbal individuals may face similar needs, depending on their condition those may vary. The fact that none of the solutions were specifically designed and conceived for aphasia individuals gives this project the motivation to try to reach an option for this population regarding their individual needs and functional profile as well as the challenges they face.

Also there are a number of custom dedicated handheld communication systems that have been developed such as Tobii Dynavox or Prentke Romich Company devices. One serious disadvantage of such systems, however, is that the combination of custom hardware and software makes the systems very expensive and out of reach for many individuals [7]. Such high prices can put these devices out of reach for many people.

Traditionally, electronic AAC devices were developed with specific hardware and software. However, due to the increasing use of mobile devices and applications, AAC software for common mobile devices, tablets, and PCs has been introduced and widely adopted. Transitioning from specialized hardware to mainstream devices can give various benefits to assistive technology users, including lower costs, increased usability, and less social stigma [3].

Talking about the mobile applications presented and the connection to the previous paragraph, price differences between customized hardware devices comparing to applications for mobile devices are clear. Also the fact that nowadays almost everyone has a mobile device and knows how to use it, can bring several advantages over customized hardware devices, such as:

- users do not need to learn how to handle a new device;
- when an AAC user faces everyday situations and needs the help of his device in order to communicate, does not have to feel different because is using a different tool from anyone else, he is just using his phone or tablet like all people do (less social stigma);

As previously stated, one of the main goals of this dissertation is to contribute to a better understanding regarding aphasia in order to study as deep as possible the challenges that people with aphasia face in their routine, thus being able to characterize the motivations and specificities of the individual with aphasia. Also, gather knowledge and explore the different assistive communication solutions that already exist for individuals with aphasia are an objective. With both of the previous goals achieved, solutions that can assist these individuals in their day-to-day communication needs may be proposed.

To achieve that, and allied to the gathered knowledge and understanding of the target users as well as the challenges they face and of a proper envisionment of the problem, methodologies to bring the user's needs and motivations to the process should be used, which in this case is a User-Centered Design approach.

After that, some primary touchpoints on the types of technology to be used according to the previous mentioned points will be defined in subsection 2.5.2, Supporting Development for Mobile Approaches.

2.5.1 User-Centered Design

User-Centered Design (UCD) is a design approach that relies on the active participation of users for better understanding of user and task requirements, as well as design and evaluation iteration. It is often regarded as the key to product usefulness and usability - an efficient method of overcoming the limitations of traditional system-centered design [17].

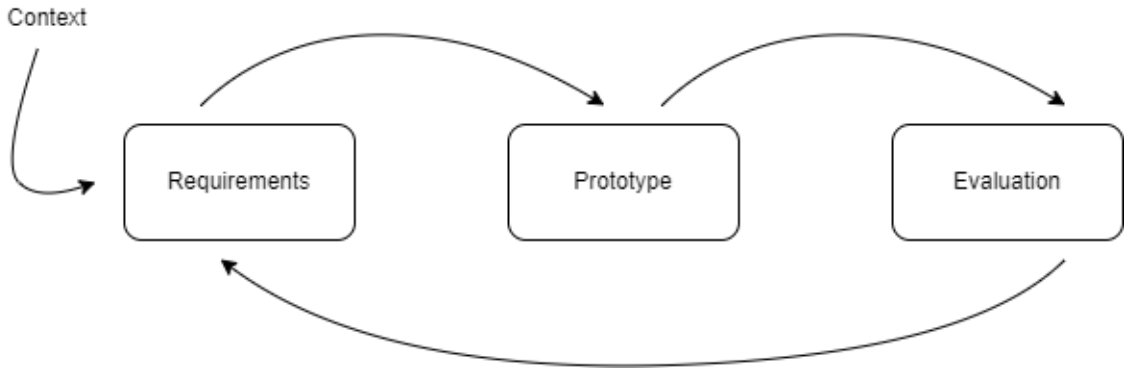


Figure 2.3: Diagram depicting the different steps for an Iterative User-Centered Design Approach. Starting from Requirements this methodology entails different short-steps of prototyping followed by evaluation which intern leads to define requirements and a novel stage of prototyping.

Therefore, UCD includes users at all user-centered design steps, and when the users may be difficult to reach, people that know them closely can be involved at early stages. Making a User-Centered Design approach Iterative is achieved by testing, failing and trying again. This process develops an understanding of the user's needs through a combination of Prototyping and Evaluation meeting certain Requirements.

Requirements

Taking into account the context of the specific situation, in this case being Aphasia and people who live with that condition in need for an AAC solution, research has to be done. This research will result in the characterization of Personas, Scenarios and Requirements that satisfies them. The Requirements Definition process often comprises steps like Obtain a clear envisionment of the problem, Brainstorming, Identifying persona expectations, Constructing context scenarios and Identifying requirements [18]. These steps represent an iterative process and designers can expect to cycle through them several times until the requirements are stable, as depicted in Figure 2.3.

Regarding **Obtain a clear envisionment of the problem**, the purpose of the design initiative is defined by the problem, that is, the problem should reflect a situation that needs changing, for both the personas and for the project, causing that user's needs become project goals. The substance of both the problem and vision statements should be derived directly from user models and research [18]. The primary purpose of **Brainstorming** is to eliminate as much preconception as possible, allowing the design to be open-minded and flexible in order to give space to imagination. Brainstorming should be unconstrained and uncritical, and too much time should not be spent on this step [18].

When talking about **Identifying persona expectations**, a question must be first answered: What is a Persona?

"Personas are user models that are represented as specific, individual human beings. They are not actual people but are synthesized directly from observations of real people" [18]. (A. Cooper, R. Reimann, and D. Cronin, 2007, p. 81)

Personas provide a precise way of how users behave, how they think, the challenges they face, their motivations, what they wish to accomplish, and why. They are based on real-world observation of real people's behaviors and motivations, which makes them appropriate and effective to use as user models due to their unique characteristics, having the ability to engage the empathy of the design and development towards the human target of the design [18].

In order to identify Persona expectations, some aspects on each primary and secondary persona must be identified such as attitudes, experiences, aspirations, and other social, cultural, environmental, and cognitive factors that influence the persona's expectations. Also, related to the product, expectations and desires the persona may have about the experience of using it must be identified and what the persona will expect from the product too. To identify all of that, some actions must be taken such as Contextual Interviews and Inquiry, Interviews with users outside of their use contexts, Information about users supplied by experts, family or fiends and Research data such as focus groups and surveys or even Data gathered from literature reviews and previous studies [18].

Context Scenarios are created before any design is performed and are used to explore how the product can best serve the needs of the personas. They should not go into specifics about the product or the interaction, as they are supposed to be broad and relatively shallow in

scope, but rather be written from the perspective of the persona, focused on human activities, perceptions, desires, motivations and mental model. **Constructing context scenarios** entails determining the primary points of contact that each primary and secondary persona has with the system, as well as possibly with other personas, over a reasonable period of time. Context scenarios should be large in reach but limited in depth [18].

After an initial draft of the context scenario is accomplished, **Identifying requirements** is possible. Requirements are often defined as "feature" or "function", and although there is a relationship between requirements and functions, it would be more appropriate to set requirements as synonymous with needs. These requirements can be thought of as consisting of objects, actions, and contexts in order to best define the human and project needs that the product must satisfy [18].

After the five steps of the requirements definition are executed as many times as wanted or needed, and successful and robust Personas, Scenarios and Requirements are accomplished as wished, an initial draft of Prototyping can start.

Prototype

Prototyping can be described as the act or set of acts in which a prototype is built, tested and modified until an acceptable outcome is achieved from which the complete product can be developed. In sum, a prototype is a mockup of the solution that is supposed to be created, that is, a scale or full-sized structural model of a project built to scale for study, testing, or display [19].

When speaking of prototyping there are two approaches that must be considered: Low-fidelity prototyping and High-fidelity prototyping. Fidelity refers to how easily prototypes can be distinguished from the final product and how they can be adjusted to highlight design features [20]. **Low-fidelity prototypes** are built to depict concepts, design options, and screen layouts rather than to simulate user interaction with the system, in other words, their primary goal is to communicate, educate, and inform rather than to train or test. They are built quickly, have limited or no functionality, and are not intended to show how the application works in detail. They are fast and inexpensive, allowing for a quick preview of the product. Storyboard presentations and proof-of-concept prototypes are examples of low-fidelity prototypes [21].

On the other hand, **High-fidelity prototypes** are fully interactive and highly functional, allowing users to interact with the user interface as if it were a real product since they are very close to the final product, realistically appearing and functioning as similarly as possible to the actual product before launch, with the majority of the necessary design assets and components developed and integrated. They are not as quick and simple to create as low-fidelity prototypes and are frequently used in the later stages to test usability and identify potential issues. Their greatest advantages are providing meaningful feedback and also, being good for demonstration in front of potential investors or stakeholders [21].

Low-fidelity and High-fidelity prototypes have many advantages and also disadvantages. As written above, Low-fidelity prototypes are constructed quickly and will save resources,

are intended to communicate, educate, and inform, but not to train or test. They are useful in identifying requirements, for example. High-fidelity prototypes are fully interactive and highly functional, allowing users to test a version of the product very close to the final one. They are not as quick and simple to create as low-fidelity prototypes, but provide great and meaningful feedback [21].

Evaluation

The Evaluation phase comes after satisfactory prototypes are achieved, weather if Low-fidelity or High-fidelity, since this phase can happen as many times as intended, like demonstrated in Figure 2.3. Its main goal is to identify and focus on specific issues, discover usability problems and how they impact the overall user experience, providing quick and inexpensive feedback. To sum up, evaluation helps determine what works well and what can be improved.

According to Jakob Nielsen and Rolf Molich, there are basically four ways to evaluate a user interface: formally by some analysis technique, automatically by a computerized procedure, heuristically by simply looking at the interface and passing judgement according to ones own opinion and empirically by experiments with test users [22]. The first three are types of analytical evaluation and do not involve users.

Also according to them, Formal analysis models have not reached the stage where they can be generally applied in real software development projects and Automatic evaluation is completely infeasible except for a few very primitive checks [22]. Having that in consideration, two approaches are left, being them Heuristic evaluation and Empirical evaluation.

Heuristic evaluation is a method for evaluating user interfaces in order to find their usability problems, that is, a group of evaluators inspects the interface against a limited set of pretty broad usability rules known as "heuristics" [23].

Regarding the previous paragraph, **Heuristic evaluation** can be understood as a method commonly used to find usability problems at different development stages of a product, that is, a process where certain rules (heuristics or guidelines that are relevant for the project) are used, by evaluators chosen by the developing team of the product, to measure the usability of the system and report issues. As heuristic evaluation is a light-weight process that can be cheap, fast, and easy to apply, it is used a lot at an early stage of the evaluation phase [24].

Regarding Heuristics, there are many alternatives that can be used, one of them being Jakob Nielsen's Heuristics. Nielsen developed the heuristics based on work together with Rolf Molich in 1990. The final set of heuristics that are still used today (among the most used and popular ones) were released by Nielsen in 1994 and establish principles such as *Visibility of system status*, *Error prevention* or *Flexibility and efficiency of use*.

Another approach that can add to the evaluation phase is **Empirical evaluation**, which uses data derived from actual observation or experimentation of users and can be also related to User-based evaluation since the last is evaluation that involves the people for whom the system is intended: users [25]. In sum, Empirical evaluation and User-based evaluation both refer to the evaluation of a theory by observation in experiments and can be used at any stage of the evaluation phase. User-based evaluation techniques or experiments can be performed

through user studies (research that focuses on understanding user behaviors, needs, and motivations) and include: Observational Methods, Focus Groups, Questionnaires such as System Usability Scale (SUS) and Post-Study System Usability Questionnaire (PSSUQ), that are post task questionnaire, Feedback from experts or from people related to the subject, among others [25].

Observation is a qualitative research method or technique where participant's behavior is observed in a natural situation. A good reason for employing **Observational Methods** is to ensure if what people say is what they actually do [26] and an example of an observational technique can be the Think Aloud Protocol where users or participants are asked to put into words what they are looking at, thinking, doing, and feeling as they complete a specific task. Observers are asked to take notes of what participants say and do, noting places where they encounter difficulty, without interfering in their actions and not attempting to interpret what they say. Another example of an observational technique is Task analysis, which is the process of learning about ordinary users by observing them performing a task and analysing how the task was accomplished regarding task duration, frequency, complexity, surrounding conditions and any other factors involved in or required to perform the given task.

Focus Groups are a form of group interview or a moderated discussion with a group of users (the group consists of a small number of carefully selected people) in which, instead of each person being asked a question, a given topic is discussed and people are encouraged to talk to one another by asking questions, commenting on each other's points of view and connecting to each other's experiences, thus enabling the gather of detailed information about a user's attitudes, desires, ideas and experiences [27]. A particular disadvantage of a focus group is the possibility of not all participants feeling comfortable expressing their honest and personal opinions in front of others and the fact that some participants may hesitate to express their thoughts when they collide with the views of another participant or when they think that it will not be well accepted.

A questionnaire is a research instrument featuring a series of questions, open and/or closed and written or oral, in order to collect useful information from respondents and can be carried out by call, video call, face to face, email or post. **Questionnaires** provide a relatively cheap, fast and efficient way of obtaining large amounts of information from a sample of people. However, a problem with questionnaires is that respondents may bend the truth as most people want to present a good image of themselves or may answer how they think society would find most acceptable.

Two types of standardized questionnaires are System Usability Scale (SUS) and Post-Study System Usability Questionnaire (PSSUQ). **The System Usability Scale** is a widely used standardized questionnaire for the quantification of perceived usability, allowing the evaluation of a wide variety of services or products such as any kind of hardware, software, websites, applications, among others [28]. The **Post-Study System Usability Questionnaire** is widely used to measure users' perceived satisfaction of a service or product at the end of a study [29]. Both Questionnaires are present in Appendix A.

2.5.2 Supporting Development for Mobile Approaches

User-Centered Design methodology has the big advantage of not only including users allong the design and development, but also of being an Iterative process in which is possible to cycle through some of the steps of UCD until the requirements are stable, as previously explored.

The fact that the used methodology to approach the user as much as possible to the process of proposing solutions based on the gathered knowledge and understanding of the people with aphasia is the Iterative User-Centered Design approach raises some challenges, once the development of the solution may continuously need to adapt as the requirements evolve when being redefined. This brings the challenge of choosing technologies that allow for iterative development and also meet user needs, and that entail characteristics like modifiability and scalability.

As stated in section 2.4, mobile solutions may have many advantages from the fact that nowadays almost everyone has a mobile device and knows how to use it to the less social stigma when an AAC user faces everyday situations and needs the help of his device in order to communicate, once he doesn't have to feel different because is using a different tool from anyone else.

To develop a native mobile application, which means creating a mobile application that is tailored and dedicated to a specified platform, that allows the iterative development of the solution there are options that make it possible, like Android or iOS.

Software development for Android and iOS requires different approaches, once that Kotlin or Java are used to develop Android mobile apps and Swift or Objective-C may be used to create iPhone applications. Android Studio and Xcode are Integrated Development Environments (IDE's) for developing Android and iOS apps respectively. Frameworks like React Native, Ionic and Flutter are for developing apps that run natively on both Android and iOS devices, also called cross-platform mobile development.

Once the goal is to achieve a solution that is not closed thus being dynamic and scalable, an approach based on services can bring many advantages like separation of concerns between the different services as well as simultaneous and autonomous development and also reusing code at scale, as services represent self-contained code packages they enable code reuse on a macro level, which may significantly reduce development time for new functionality. Also, the fact that it allows the integration of features available in the cloud through, for example, REST API or API once they induce desirable properties, such as performance, scalability, and modifiability.

2.6 Conclusions

Taking into account everything that has been said in this chapter, from the importance of communication, to in what consists properly aphasia and solutions provided by Augmentative and Alternative Communication devices or applications, the reached conclusion is that solutions need to be proposed in order to try to provide people with aphasia with a specific

solution to their specific needs, based on AAC, that aims to keep these people socializing, maintaining a job, attending school, or living as independently as possible. To achieve that, the characterization of Personas, Scenarios and Requirements that satisfies them has to be first done.

CHAPTER 3

Personas, Scenarios, Requirements

In this chapter, the methodologies used to achieve the Personas, Scenarios and Requirements will be presented, as well as all the Personas, Scenarios and Requirements. Aphasia individuals were modeled into Personas, how the system may help them tackle some challenges in particular contexts into Scenarios and a first list of Requirements was identified based on actions present in the accomplished context Scenarios.

3.1 Methods

When a clear envisionment of the problem was being obtained, and in conversation with experts in this matter and while brainstorming, it became perceivable that people diagnosed with any kind of Fluent Aphasia would likely to take advantage of what was being idealized, and because of that it was decided to start an approach for these people.

By using personas, an understanding of users' goals and needs in specific contexts can be developed. Taking into account the context, that in this case is focused on people diagnosed with Fluent Aphasia, the research to achieve such Personas was not only based on data collected from literature and studies, but on information provided by experts, such as Speech and Language Therapists and also on research data such as a focus groups carried out with a Human Computer Interaction (HCI) Expert, a Speech and Language Therapist (SLT) and other elements of the project in which this document is inserted in.

In this particular context, the Scenarios are descriptions of Personas using the system in order to achieve a goal or perform one or more tasks. Such narrative was based on points of view that try to be representative of specific situations that people diagnosed with Fluent Aphasia may face, taking into account, once again, literature and studies, information provided by experts and also on research data such as focus groups.

After Context Scenarios and Personas are on a solid base, identifying requirements is the next step. Requirements were extracted from the objects, actions, or contexts depicted in the scenarios in order to best determine what needs, goals and motivations of the user the product must satisfy.

3.2 Personas

Personas are based on research in order to represent the different user types that might make use of the product, and taking into account such research, four Primary Personas and four Secondary Personas were achieved. Primary Personas are representative of four people with Fluent Aphasia, namely Wernicke Aphasia, Conduction Aphasia, Transcortical Sensory Aphasia and Anomia Aphasia. Each one of the primary Personas presents different types of characteristics, needs and challenges regarding aphasia having in consideration their own individual personality but also the differences between the different types of Fluent Aphasia. The Secondary Personas that are presented are related in some way to some of the Primary Personas, and are intended to highlight and contextualize some of the characteristics of the related Primary Personas.

- **Primary Persona:** Wernicke's Aphasia Man
- **Name:** Manuel Almeida
- **Age:** 56 years old
- **Profession:** Accountant

Table 3.1: Main characteristics of Manuel, a Persona with Wernicke Aphasia. Image from *Pexels* by Kampus Production.

Manuel Almeida is 56 years old, having been born in Águeda, a city in the district of Aveiro, on December 25, 1965. He is married and has three underage children, ages 16, 14, and 9, all of them are boys. He lives in Águeda with his wife and children.

Manuel worked as an accountant and dealt with a lot of stress, mostly in the end and beginning of each month. The fact that Manuel has three kids at school also gave him great amounts of stress, having to deal with, for example, three different schedules and pick up times, and help the kids with homeworks. This frenetic lifestyle gave him little time to focus on himself and he didn't take much time to have a balanced diet regarding nutrition and sports, being overweight.

He is very familiar with the use of smartphones and technology in general because he used to spend most of his working day on the computer and also made use of some mobile applications to help him organize his kids schedules and specificities.

Due to his very stressful and not very healthy way of living, Manuel had a stroke last month, which led him to lose some of his communication skills. His reading and writing abilities have been harmed, and he also has difficulties with naming, repeating, and comprehending oral language, in other words, he has been diagnosed with Wernicke Aphasia.

Taking into consideration Manuel's recent condition, not only his job is compromised but also his way of living in general. He is not working since the brain injury episode and he is not taking care of his kids schedules and helping on their homeworks anymore, as his communication difficulties won't let him perform his work as before and he also feels emotionally affected.

Another reason is that, since the brain injury episode, Manuel is attending a lot of appointments, such as speech and language therapy, physiotherapy, occupational therapy, neurology and cardiology, as doctors are still studying the possible causes of the stroke, trying to anticipate future episodes and to better his condition, giving him little time to focus on anything else but his health condition.

Manuel's brain injury episode was really a wake up call for him and he does not want to go back to his old lifestyle. Manuel wishes to improve his health condition and stop being overweight, by practicing exercise and start to pay attention to his food choices, being then able to be more present in his kids' lives. To accomplish that, Manuel started to attend the gym, but most of the times he struggled when the staff approached him trying to help with the gym machines or when someone just wanted to make some conversation and that led Manuel to just leave early.

Since Manuel loves his work and he is very focused on what he does, his greatest obstacle fin this aspect is the fact that he struggles with comprehending oral language, has difficulties with reading and writing and also with performing calculations, aspects which can compromise a lot his communication with colleagues and clients and also his performance as an accountant.

Motivation: Manuel's greatest motivations are to feel confident regarding his communication with others, being able to express what he aims and understand what others say, for example when going to the gym, at work or with his family, whether to ask them to speak slower or tell them his opinion.

- **Secondary Persona:** Mother of Manuel, the Man with Wernicke Aphasia
- **Name:** Maria Almeida
- **Age:** 81 years old
- **Profession:** Retired

Table 3.2: Main characteristics of Maria, Mother of Manuel who is the Persona with Wernicke Aphasia. Image from *Pexels* by Kampus Production.

Maria is 81 years old and she was born in Águeda where she lives until the present day,

on October 26, 1940. Maria has two children, one of them being Manuel, the man diagnosed with Wernicke Aphasia.

Since Maria became a widow, when her husband passed away in 2017, Manuel has been a great support to her. They used to talk everyday on the phone and on the first weekend of every month Maria used to go to her son's house to spend some time with him and her grandchildren.

When Manuel had a stroke, last month, losing some of his communication skills and was later diagnosed with Wernicke Aphasia, Maria became really worried about him due to his health condition and also sad because Manuel doesn't talk to her everyday like before. And, when the opportunity arises he often doesn't understand what Maria says because she gets too enthusiastic and speaks very quickly or keeps changing subjects. Or Maria doesn't understand what Manuel says, due to his difficulties with oral expression.

Maria totally understands that Manuel doesn't feel comfortable talking to her everyday, but it is like she lost that one hour of company, when she would talk about her day, ask about Manuel's day and also how the kids were doing. He always had the capability of making her smile, and she feels down for not being able to do the same and doesn't know what to do to make it easier when both of them are communicating or to make him comfortable.

Motivation: Maria's greatest motivation is that she wants to talk to Manuel everyday like before, and in order to do that she wants to learn more about his condition and how to properly communicate with her son.

- **Primary Persona:** Conduction's Aphasia Woman
- **Name:** Ana Tavares
- **Age:** 52 years old
- **Profession:** CEO

Table 3.3: Main characteristics of Ana, a Persona with Conduction Aphasia. Image from *Pexels* by LinkedIn Sales Navigator.

Ana is 52 years old, she was born in Porto, on March 26, 1970. She is divorced, has no kids and lives by herself in Porto.

Ana worked as CEO of a cosmetics company also based in Porto. She was very focused on her career and very passionate and proud of the company she built from scratch.

Ana spent most of her day at work, whether in meetings, business trips or business dinners. She even used to answer calls on the weekends and never had a no as an answer to work

related commitments or subjects.

One of the things that helped Ana to overcome the stress and busy schedule she had, were cigarettes. She has smoked since 20 years old, when attending business university, and even though she tried to quit smoking a few times, she never succeeded and always failed when the stress would hit.

Ana is very familiar with the use of technology, as she not only was used to make video calls related to work all the time and made use of applications to organize her schedule, she also has an online platform of her company in which she took great power of decision and choice in the development, taking into account her experience and vision regarding similar platforms. Also, as Ana was very busy and had little time to spend in supermarkets and malls, all her shopping was done online.

Being a smoker and having a stressful lifestyle led Ana to suffer from a stroke recently. She has been diagnosed with Conduction Aphasia, which harmed some of her communication skills, such as her ability of repeating, naming, reading and writing.

Since the brain injury episode, Ana is not working and feels very depressed, because her company is and has always been her biggest motivation and reason for living. She tried to go back to work last week, attending some business meetings, which led her to be even more depressed due to the fact that she felt that she was being put aside due to her condition, not being asked as many questions and opinions on decisions as before. Ana has always been seen and feels like a powerful businesswoman, and to see that status compromised puts her down a lot.

Motivation: Ana's main motivation is to go back to work and to be able to perform it exquisitely as before, giving her opinion and point of view everytime she feels like it, that is, expressing herself with confidence again and not being put aside.

- **Secondary Persona:** Speech and Language Therapist of Ana, the Woman with Conduction Aphasia
- **Name:** Romeu Silva
- **Age:** 42 years old
- **Profession:** Speech and Language Therapist

Table 3.4: Main characteristics of Romeu, Speech and Language Therapist of Ana who is the Persona with Conduction Aphasia. Image from *Pexels* by cottonbro.

Romeu Silva is 42 years old and he is from Coimbra. He was born on February 16, 1980. After attending Aveiro University and finishing his degree on Speech and Language Therapy,

he headed to Lisbon looking for job opportunities, where he remains to this day. He is single and lives by himself.

After some years in Lisbon, Romeu took a master degree in Speech and Language Therapy specialized in Neurological Disorders in Adults in the University of Lisbon, and that's how Ana heard about him, because she has a friend that was a teacher of Romeu and recommended him.

Since he started being Ana's therapist, he really empathizes with her, maybe because their age is not that different and their goals of prospering professionally are similar, which leads Romeu to understand what she must be feeling. However, he knows it will take some time until she regains her natural speech again, something that is crucial in order she can go back to work without insecurities.

Motivation: Romeu's main motivation regarding Ana is that she can overcome as soon as possible her communication difficulties. In that sense, something that would help Ana to be motivated, to practice and to face her fears outside their sessions would help Romeu and Ana to speed up the process of recovery.

- **Primary Persona:** Transcortical Sensory's Aphasia Man
- **Name:** Joaquim Barbosa
- **Age:** 78 years old
- **Profession:** Retired

Table 3.5: Main characteristics of Joaquim, a Persona with Transcortical Sensory Aphasia. Image from *Pexels* by Steshka Willems.

Joaquim Barbosa is 78 years old, having been born in Viseu, where he lives until this day, on February 18, 1944. Joaquim lives with his wife, Maria, and even though their son Duarte no longer lives with them, they are visited weekly by him, his wife and their son on Sundays for lunch.

In his working days, Joaquim was a bus driver in his city. He used to love his profession, as he made a lot of good friendships with colleagues and passengers because he has always been a very communicative person.

Since Joaquim retired, and because he always was a very active person, he did not want to stop and just be at home. He and his wife decided to adopt a dog at the Viseu animal shelter, Luna, who not only kept them active at home with her crazy games but also outside the house, on the daily walks they both took with her. Joaquim and his wife are also responsible

for taking care of their grandchild, Martim, every Wednesday afternoon, as he has a free afternoon and his parents are at work. Having Luna and Martim playing at their house and spending some quality time with Martim, like listening to him telling all the crazy things he did the previous week with his friends at preschool, means everything to Joaquim.

As Joaquim has a very close group of friends and they are all retired, they decided to meet for dinner every Saturday and spend a great time around the table together, since all the excuses related to work and kids no longer exist.

Joaquim is not an expert regarding technologies, but he is an easy learner. Since he retired he has a smartphone and makes use of some mobile applications like Facebook, Messenger and Youtube.

Three months ago Joaquim had a stroke, which not only paralysed his right side arm, but left him with Transcortical Sensory's Aphasia that affects his capability of communicating normally, as he struggles with comprehending oral language and also with repeating, naming, reading and writing.

Since then, Joaquim can no longer go for walks with Luna alone, and barely can play with her or Martim. The most heartbreaking part for Joaquim is that he cannot understand most of the things Martim says when telling about his week at preschool or about the new jokes he learned. He also stopped attending Saturday night diners because he doesn't feel confident going there due to his communication difficulties.

Joaquim is really trying his best to get better in physiotherapy and speech and language therapy, but as the doctors said, it will be a long process.

Motivation: Joaquim aims to take the most advantage of speech and language therapy in order to better his fluency and thus be able to communicate again with Martim and hear all the amazing things he has to say, as before. Being with his friends again, without feeling inferior, is also a motivation to Joaquim.

- **Primary Persona:** Anomia's Aphasia Woman
- **Name:** Judite Rodrigues
- **Age:** 38 years old
- **Profession:** Teacher

Table 3.6: Main characteristics of Judite, a Persona with Anomia Aphasia. Image from *Pexels* by Christina Morillo.

Judite Rodrigues is 38 years old, having been born in Lisbon, on January 17, 1984. She

is single, has a daughter, Inês, and lives with her parents, both of them retired, and Inês in Lisbon.

As Judite never had siblings and always wanted to, being a mother of Inês and working with children was a dream come true. Judite worked as a primary school teacher and loved what she did. She is a very caring person, and in addition to the dedication to Inês and her students, she also took great care of her parents. She worked to provide them the best life possible, as they once did towards her.

Since Judite was 11 years old, that she lives with type 1 Diabetes, for which there is no cure. Regarding her condition, it is very important to keep the diabetes stable and controlled and to achieve that, Judite needs daily insulin injections in addition to other preventive measures, including the use of medication, healthy eating and physical exercise.

In addition to taking advantage of mobile applications to help her organize her schedule and Inês's, Judite also made use of her computer to prepare lessons, as she is very into finding dynamic and thematic approaches of teaching, for which she is well known in the teaching community, as she used to share some tips on social networks and has great feedback.

Since Inês started attending primary school last year, Judite had been under more pressure as her work load and responsibilities increased a lot. In addition to having work related to school to do (correct tests and assignments, prepare lessons, etc) and also helping her mother (with everything related to the maintenance of the house, grocery shopping and meal preparation so as not to overwhelm her), she also helped Inês with her homework and school tasks, and since then, from the moment she got home, she never stopped.

Because of all that, Judite stopped exercising and taking care of her healthy eating habits as she used to. Judite had a stroke two months ago, which led her to be diagnosed with Anomic Aphasia. She now struggles with naming and in addition to that, suffers from paralysis on the right side of the body.

Since the brain injury episode, Judite struggles with helping at home like she used to. Seeing her parents taking care not only of Inês, but also of her and the house, brings sadness to Judite as she feels that she is not providing the life she wanted for them. She tries to help Inês with her school related tasks, but not as effectively as before, because she can not always find the right words and even though she tries to get around it by explaining what she means, it's not the same thing.

Judite is not working, once the naming barrier can strongly compromise her teaching, and even though, at home, she can always try to explain what she means when the words fail, at school is not the same thing. Judite is now attending speech and language therapy and physiotherapy appointments weekly in order to work towards regaining her full communication and physical capabilities, but the doctors have already told her that it will take some time until then.

Motivation: Judite's greatest motivation is to go back to work and to her students as soon as possible, without the naming barrier and the paralysis interfering. Being able to help Inês again with her homeworks like before and do all the things she used to do outside the

house without feeling anxiety are also goals for Judite.

- **Secondary Persona:** Mother of Judite, the Woman with Anomia Aphasia
- **Name:** Natália Rodrigues
- **Age:** 68 years old
- **Profession:** Retired

Table 3.7: Main characteristics of Natália, Mother of Judite who is the Persona with Anomia Aphasia. Image from *Pexels* by RODNAE Productions.

Natália Rodrigues is 68 years old and was born on April 13, 1954. Natália is from Mozambique and when she was 13 years old she moved to Lisbon, where she lives until the present day with her husband, her granddaughter and her only daughter, Judite, the person diagnosed with Anomic Aphasia.

Before retirement, Natalia was a nanny in a nursery, but she decided to retire when she was 64 years old because, even though she loved her job and kids, it was time to rest and enjoy life. Since then, she was having a great life with her husband, as they were attending outdoor gymnastic classes and would meet with friends in the neighborhood almost every day, until everything changed when, two months ago, her daughter had a stroke.

Natália and her husband stopped attending dance classes and they only get to see their friends when they visit, as Natália almost never leaves the house because she has no time considering that Judite and Inês need her around almost all day.

Having Judite with Anomia Aphasia and suffering from paralysis on the right side of her body breaks Natália heart, and even though Natália doesn't feel the naming deficit that Judite faces, as Judite is comfortable at home taking her time explaining what she intends to say, Natália knows that in order to go back to work, Judite needs to overcome that in some way. She also has been aware that Judite is sometimes apprehensive about leaving the house to do things she used to, afraid of being asked something she may struggle with when responding.

Motivation: Natália's greatest motivation is to see her daughter happy again, and in order to that, she knows Judite has to be back to work and be capable of facing some of her fears regarding her condition. She also aims to learn more strategies that can help Judite regarding her condition. With Judite on the right track, Natália may think about going back to dance classes and get to see her friends more often.

- **Secondary Persona:** Daughter of Judite, the Woman with Anomia Aphasia
- **Name:** Inês Rodrigues
- **Age:** 7 years old
- **Profession:** Primary School Student

Table 3.8: Main characteristics of Inês, the Daughter of Judite who is the Persona with Anomia Aphasia. Image from *Pexels* by Monstera.

Inês Rodrigues is 7 years old and was born on 14 May, 2015, in Lisbon. She lives with her mother and her grandparents, also in Lisbon. She is attending the second grade at the primary school, where her mother, Judite, teaches, and she is a great student.

As soon as Inês entered school, instantly became super excited about it. The fun games, the weekly picnics during school playground, the amazing birthday parties everyone got, made all of a sudden, learning and going to school fun.

When, two months ago, Judite had a stroke and stopped attending school, Inês became really sad, as Inês and Judite always went to school together and always came from there together and Inês loved that. Since then, her grandmother, Natália is the one to take her and pick her up from school. Also, Inês's grades went down, and it's not that her teacher is bad, but for Inês is not the same thing as it was when her and her mother were going together to school and when her mother helped her so well with the homeworks.

Inês also misses talking to her mom almost until falling asleep as she used to, telling her about her day at school, her insecurities regarding friends or even listening to mom's stories when she was herself a student.

Motivation: Having her mother to help her again with the homeworks like she used to and having her mother to go with her to school are some of Inês's motivations, but the greatest of all motivations is to confide everything to her mom as before.

Scenarios are representative of daily situations that the previous Primary personas may face, whether they have been just diagnosed, in recovering or already recovered, making use of the system. To accomplish the following Scenarios, Personas motivations were at the centre and having that in consideration, nine context scenarios were developed. They were validated by Speech and Language Therapists with a strong experience with patients with aphasia.

3.3.1 Manuel's Context Scenarios

In the following context scenario, Manuel, one of the Primary Personas, is establishing a phone call with his mother, who now is prepared with dialogue strategies taking into account Manuel's condition, and through the use of an application manages to overcome some of his communication difficulties regarding his diagnosis with Wernicke Aphasia.

- **Manuel talks to his mother on the phone:** Manuel is at home, helping his wife to prepare dinner, when his home phone rings. It is his mother. He answers the call and puts it on speakerphone. Manuel and his mom start talking and everytime a new topic of conversation is introduced by Maria, she clearly identifies it: "I'm talking about the kids." As soon as Manuel's first difficulty understanding what his mother is saying arises (Oral Language Comprehension difficulty), he looks into the application and a display of images and words describing what his mom just said are depicted. Manuel manages to understand what he previously didn't and the conversation keeps going.

In the next context scenario, Manuel, one of the Primary Personas, is at the gym. When a staff member approaches trying to help him with the machine, Manuel manages to overcome some of his communication difficulties, regarding his diagnosis with Wernicke Aphasia, through the use of an application.

- **Manuel goes to the gym:** Manuel is at the gym and after doing some cardio, decides to try a chest press machine. While Manuel struggles on getting the right way to use it, a personal trainer approaches him in order to help Manuel on using it the right way. As soon as he starts talking to Manuel, he gets a bit nervous and doesn't understand what the personal trainer is saying (Oral Language Comprehension difficulty). He then presses the application button that issues the following message: "I'm Manuel, and due to my most recent diagnosis of a post-stroke Language Impairment, I had difficulties in understanding what you just said, could you repeat it a little slower, please?". The personal trainer repeated it slower and Manuel, who after that became more relaxed, could keep a circumstance conversation with the personal trainer and was able to use the machine properly.

In the next context scenario, Manuel, one of the Primary Personas, is back at work. When talking to a client, Manuel manages to overcome some of his communication difficulties, regarding his diagnosis with Wernicke Aphasia, through the use of an application.

- **Manuel returns to work:** Manuel is at work when a co-worker informs him that a client wants to speak with him. Manuel receives the client and soon realizes that he has doubts about filling in the IRS, which Manuel wills to clarify. When Manuel starts to explain what he intends, a difficulty finding a specific word arises (Naming difficulty), and Manuel, while trying to explain it in other words, looks to the suggested words and images by the application. As soon as he manages to find the correct word, the explanation goes on.

3.3.2 Ana's Context Scenarios

In the next context scenario, Ana, one of the Primary Personas, is back at work, and through the use of an application, manages to overcome some of her communication difficulties and is thus able to capacitate a co worker regarding her diagnosis with Conduction Aphasia.

- **Ana prepares a meeting at work:** Ana is at work, and as she is aware that some people, knowing her condition, sometimes put her aside (not asking her as many opinions as before or not interacting as much with her, simply because they don't know how to deal with communication difficulties who might happen) decides to prepare an important meeting that will occur the next day with Pedro, her right-hand man, who will be conducting the meeting. As soon as they start, Ana puts the application on listening to their conversation and as it goes on, and when Pedro starts to speak a bit fast, the application reproduces the phrase "You are speaking too fast, maybe try to slow down!" or when Pedro starts to build huge sentences with many concepts mixed in, the application reproduces the phrase "You are building too many long and complex sentences, try to use simple sentences", which not only allows Ana to follow the dialogue better but also enables Pedro to better learn how to help Ana, that is, capacitates him regarding her communication needs. As Pedro becomes more aware of what Ana faces on a daily basis, asks her if there are some communication guidelines he should be informed of, to which Ana shows him a topic available on the application: "Rules of good communication".

In the next context scenario, Ana, one of the Primary Personas, who is back at work, with the help of her speech and language therapist, Romeu, and through the use of an application, manages to overcome some of her communication difficulties regarding her diagnosis with Conduction Aphasia.

- **Ana attends a meeting at work:** Knowing the importance of the meeting that will occur, Ana previously scheduled an appointment to the end of the previous day with Romeu, her therapist, in order to prepare some topics she wants to address. She then prepares her speech with Romeu, which advises her to record some key sentences on the application. The next day, and during the meeting, Ana was able to see a big difference regarding Pedro's approach towards her, asking her point of view much more often,

speaking in her direction a lot more and asking other participants of the meeting to slow down in some particular moments. When the time for Ana's speech arrived, she was able to say what she intended in most of her approach, as her early practice gave her some confidence in that matter, and just in the end of it, and maybe because of tiredness, she struggled in a sentence and made the application reproduce it.

3.3.3 Joaquim's Context Scenarios

In the next context scenario, Joaquim, one of the Primary Personas, is having his grandson at the house, because it is Wednesday, and through the use of an application and when telling Martim a story, manages to overcome some barriers brought by his diagnosis with Transcortical Sensory Aphasia.

- **Joaquim tells his grandson a story:** Joaquim is at home and Martim is about to arrive to spend the afternoon. Joaquim, knowing his struggles with reading (Reading difficulty), decided to introduce a little story that he wishes to tell to Martim on the application, with the help of his wife. After Martim arrived and when Joaquim was ready to tell him the story like he used to, puts the story on display on his headphones, and repeated it to Martim. Even though Joaquim didn't understood most of what he was saying due to his Oral Language Comprehension difficulty, he was still happy to be telling a story to Martim as in the old days.

In the next context scenario, Joaquim, one of the Primary Personas, is taking a walk with Luna and remembers he needs to get food for her at the supermarket. Through the use of an application, manages to overcome some barriers regarding his diagnosis with Transcortical Sensory Aphasia.

- **Joaquim goes to the supermarket:** Joaquim is at home, and with the knowledge and approval of his physical therapist, decides to go for a walk with Luna. Before Joaquim leaves the house, his wife asked him if he could get some food for Luna at the supermarket, to which Joaquim said yes, and before leaving the house, took a picture of the packaging and saved it in the application Favorites interface. Joaquim then takes the walk with Luna and after that reaches the supermarket. As he is having difficulties finding the section of the supermarket which has dog food realizes he needs to approach an employee to help him find it. He then approaches the employee he found first, and asks him for help. The employee asks him which brand is the food, and Joaquim doesn't understand what the employee said (Oral Language Comprehension difficulty) and gets nervous. He then presses the application button that issues the following message: "I'm Joaquim, and due to my most recent diagnosis I had difficulties in understanding what you just said, can you be patient please?". The employee nodded and waited for Joaquim to show him the picture of the package of Luna's food. The employee then showed Joaquim the way to get to that section and Joaquim was able to find the right food for Luna.

3.3.4 Judite's Context Scenarios

In the following context scenario, Judite, one of the Primary Personas, is helping her daughter to do homework, and through the use of an application, manages to overcome some of her communication difficulties regarding her diagnosis with Anomia Aphasia.

- **Judite helps her daughter with her homework:** Judite is at home, helping Inês to do homework, and faces a difficulty finding a specific word (Naming difficulty). Judite then draws on the app white board what she intends to say and before making the application guess it shows it to Inês. Inês tries to guess what her mom intends to say, until Judite noded with her head when Inês said the right word. They keep going with the homework and it is kinda fun for Inês when this happens because it seems like a mimic game.

In the following context scenario, Judite, one of the Primary Personas, is back at school, and through the use of an application previously installed on her smartphone, manages to overcome some of her communication difficulties regarding her diagnosis with Anomia Aphasia.

- **Judite is back to being a teacher:** Judite is at school, giving a support class to some students with difficulties, and struggles in finding a specific word (Naming difficulty) when preparing to explain a topic to her students. She knows the word starts with a "p" and that is a fruit, and decides to get into the application, goes to the topic "Fruits" and types the first letter of what she means and suggested words and images of what she might want to say are displayed, making Judite able to find that the right word was "pêra" and then start her explanation to the kids.

3.4 Requirements

Requirements were identified based on the different actions and features present in the accomplished Scenarios for the Personas, providing the basis for designing and continuously evaluating the product itself. With that being said, the applications should allow the user to:

Priority	Requirements
1	Transcribe Speech in real time. Record Voice. Hear previous recorded Voice. See transcription of recorded Voice. Save recorded Voice and transcription of it. Access saved Voice clips and respective transcription. Insert Text. Synthesize Text. Save Text and respective Synthetization. Access saved Text and Synthetized audio. Use the device camera to identify objects. Hear the identified object. Sketch. Identify sketched objects. Hear the identified sketched object. Access words associated with different contexts. Search for words. See images associated with words. Hear the selected word.
2	User authetication. Receive notifications when speech is being transcribed. Get a summarization in the end of a conversation. View summary conversation highlights. Be able to click on the conversation highlights. See images associated with the conversation highlights. Save images, texts and audios as favorites. Access images, texts and audios saved as favorites. Access rules of good communication. Access information relative to Aphasia. Ask for help.
3	Add new Context relative to Perform Task interface. See a tutorial right after entering the app. See explanatory texts at each interface. See the different people participating in a conversation properly identified. Increase or decrease the font size of the transcribed speech. See a timeline at the top of each interface. Click the timeline. Be redirected to the the page clicked on the timeline. Access rules of good communication directly from the Speak interface.

Table 3.9: Identified Requirements based on the different actions and features present in the accomplished Scenarios, divided by priority.

Table 3.9 depicts the identified Requirements based on the different actions and features

present in the accomplished Scenarios as well a priority level associated to each requirement. The priority was defined taking into account the objectives of the project, as priority level 1 requirements are considered essential ones, once they have the features that make the system useful and viable, priority level 2 complementary ones and priority level 3 beautification add-ons requirements.

Looking at the scenarios as well as the consequent requirements, it is possible to observe that the contexts in which the user makes use of the system as well as the features present in it, naturally tend towards the use of a mobile device. That allied to the many advantages that mobile devices have, already mentioned in this document, such as ease of use once that nowadays almost everyone has a mobile device and knows how to use it, portability and less social stigma associated when an AAC user faces everyday situations and needs the help of his device in order to communicate, not feeling different because is using a different tool from anyone else, has led to the decision to approach the system in the form of a mobile application.

3.5 Conclusions

Since the main goal of this dissertation is trying to achieve a solution that may be helpful for people diagnosed with Fluent types of Aphasia, and since people are the main focus of the desired product, in the current chapter people with such condition were modeled into Personas, how the system may help them tackle some challenges in particular contexts into Scenarios and their needs into Requirements. In sum, a clear understanding of the system's users was obtained.

Looking at Figure 2.3 of chapter 2, where a diagram is depicting an Iterative User-Centered Design Approach, it is now safe to say that, having in consideration the context, reliable Requirements were achieved and is now time to evolve to the next step, Prototype.

CHAPTER 4

Iterative Development and Evaluation of a Low-Fidelity Mockup

In this chapter, an acceptable low-fidelity prototype was achieved after three stages of development and evaluation took place, carried out with a Focus Group composed by a Speech and Language Therapist (SLT) and a Human Computer Interaction (HCI) Expert. All of the low-fidelity prototypes were developed with the main goal of reaching an agreement on how the concepts and some priority requirements identified in chapter 3 would be organized, and therefore, not going into too much detail on the specifics of each page.

After the Low-Fidelity Prototype was achieved, some visual guidelines regarding the system were defined with the help of a Designer with three years of working experience in application development.

4.1 Low-fidelity Prototype

Before evolving to the prototyping process, and having in consideration all the previous steps that culminated in the Requirements present in chapter 3, the overall main idea at this point was to achieve a mobile solution, designed in European Portuguese, since the goal is to work with Portuguese speaking individuals diagnosed with Fluent Aphasia, with four main interfaces identified, being them Favorites, White Board, Transcription and Camera. All of the remaining features identified at the Requirements phase are supposed to be part of one of these interfaces, giving the user few options to decide and thus providing a simple and minimalist approach. Also, it was intended to present these interfaces to the user with simple language and representative icons.

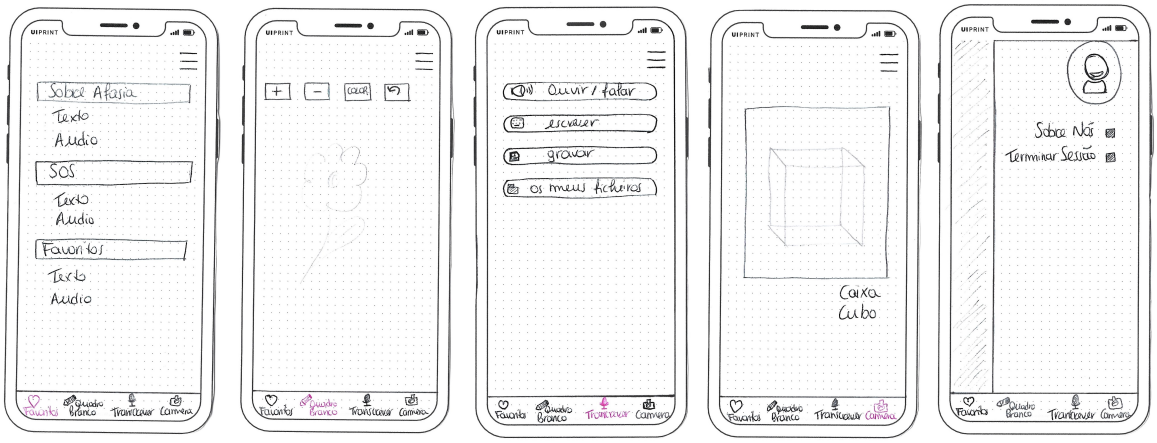
The prototype went through different iterations, and each low-fidelity prototype is an improved version of the previous one, that means that the second prototype, Figure 4.2, is an improved version of the first prototype, Figure 4.1, and by the same logic, the third prototype,

Figure 4.3, is an improved version of the previous ones. For each proposal, feedback was obtained in discussions with a Speech and Language Therapist (SLT) and a Human Computer Interaction (HCI) Expert that compose a Focus Group. Also, feedback from a Designer was used.

All of prototypes were hand drawn, given the constant and quick changes this kind of prototypes face, since they are a very primary version of what is intended, and also, since their goal are to be quick and efficient.

4.1.1 First Low-fidelity Prototype

The first achieved Low-fidelity prototype was designed having in mind a navigation bar that would contain the four main identified features of the product, being them the Favorites interface, the White Board interface, the Transcription interface and the Camera interface.



(a) Favorites page **(b)** White board page **(c)** Transcription page **(d)** Camera page **(e)** User page

Figure 4.1: First achieved Low-fidelity prototype, containing the main interfaces of the application, Favorites, White Board, Transcription, Camera and the User Page.

As is possible to see in Figure 4.1.a, the Favorites page would have pre defined important sections considered relevant to the user as well as items saved as favorites by the user. It would contain a first subsection named "About Aphasia" with a text and an audio describing aphasia for anyone related to the user who may want to know more about the condition, a second subsection named "SOS" with a text and an audio ready to be displayed if the user felt that he was in an emergency situation and a third subsection named "Favorites" with saved items from the following pages that the user felt were relevant to put in spotlight, such as images, audios and others.

The White Board page, Figure 4.1.b, would allow the user to draw what he may want to say but couldn't because of his diagnosis, allowing him to show others during a conversation for example.

The Transcription page, Figure 4.1.c, would allow the user to make use of functionalities like transcription or synthetization as well as access them if previously saved by him. It would contain four subsections, being them "Listen/Speak", allowing the user to maintain a conversation with other people and see the written transcription in realtime, the subsection

"Write", allowing the user to write what he wants to and be able to listen what he just wrote, the subsection "Record", enabling the user to record an audio he may want to display in the future and a last subsection named "My files" allowing the user to access previous recorded audio and text files.

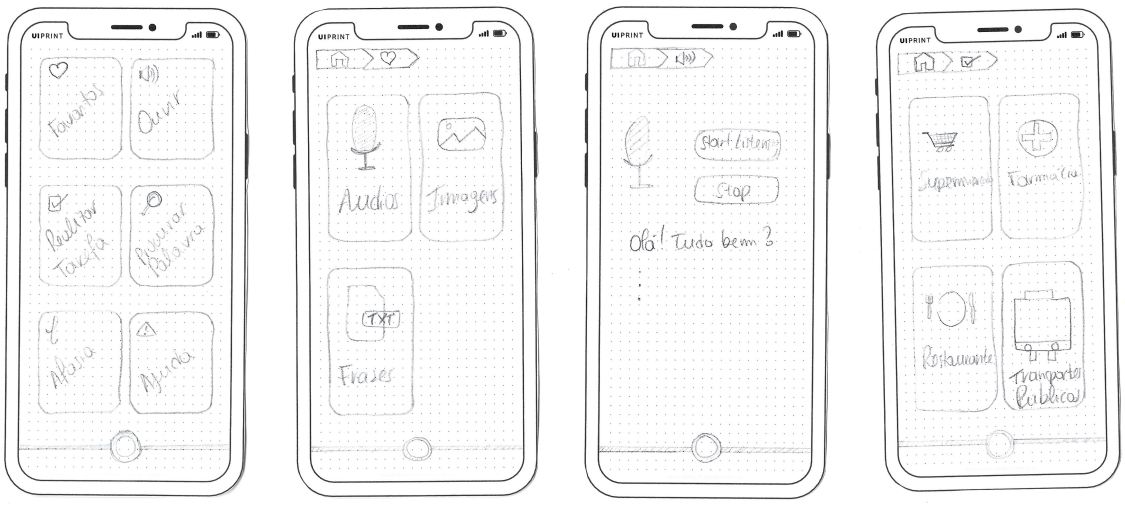
The Camera page, Figure 4.1.d, would allow the user to point the phone camera to any object and the application being able to identify it, helping the user to recognize objects or things he could not due to his condition. The User page, Figure 4.1.e would be accessible from an icon present in every interface on the top right corner and was intended to be a simple page where the user could find out more about the system or log out. It would contain a photo of the user, an "About Us" subsection containing a brief description of the product and a "Log Out" subsection for the user to log out.

After an evaluation of this Low-fidelity prototype was performed, a lot of issues were identified, being them:

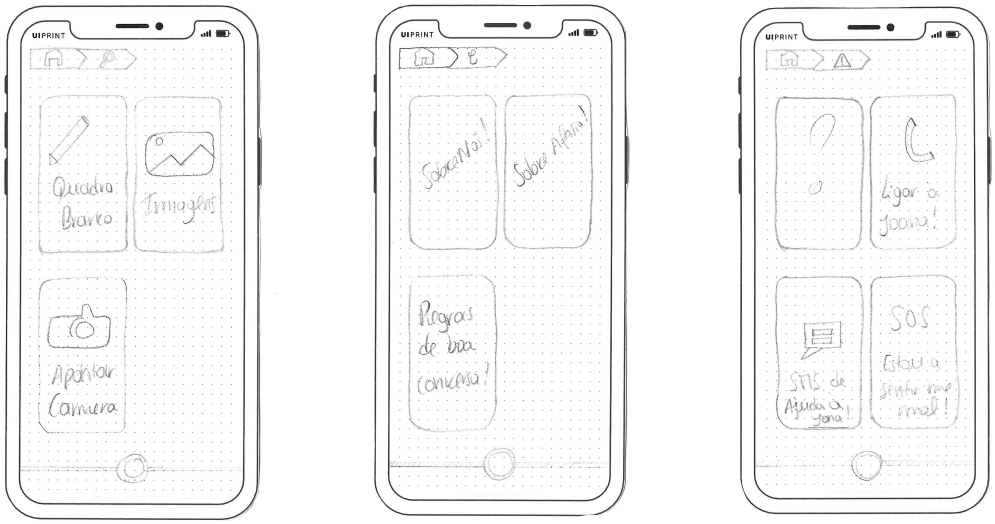
- The Navigation bar may be an issue for people with body paralysis, leading to some page changes by mistake.
- Navigation bar icons and text are too small and barely noticeable.
- "About Aphasia" and "SOS" should be accessible from the first page and as relevant as the other four identified main features, given their importance.
- On the page "Transcription" too many options for the user to choose and also very technical options.
- Adding clickable breadcrumbs or navigation timeline to the pages would allow the user not only to locate themselves in terms of the navigation of the application but also allow him to return back.

4.1.2 Second Low-fidelity Prototype

Having in consideration the Evaluation performed on the first Low-fidelity prototype and the issues that were identified, a second approach of a Low-fidelity prototype was designed and achieved.



(a) Menu page **(b)** Favorites page **(c)** Listen page **(d)** Perform Task page



(e) Search Word page **(f)** Aphasia page **(g)** Help page

Figure 4.2: Second achieved Low-fidelity prototype, containing the main interfaces of the application, Favorites, Listen, Perform Task, Search Word, Aphasia and Help.

As Figure 4.2 depicts, the navigation bar was replaced by a full-sized Menu page, Figure 4.2.a, containing six buttons with bigger icons and text, each button corresponding to the main relevant features identified, being them “Favorites”, “Listen”, “Perform Task”, “Search Word”, “Aphasia” and “Help”.

Also, an approach to the mentioned clickable breadcrumbs was attempted in the six following pages, that were also redesigned. The Favorites page, Figure 4.2.b, would contain the items selected by the user to be favorite divided by their type, that means, audio files, image files or text files.

The Listen page in Figure 4.2.c, would allow the user to directly access the feature of maintaining a conversation with other people and see the written transcription in realtime, just clicking in the button “Start Listening” to enable it and stopping by clicking in the “Stop Listening” button.

The Perform Task page in Figure 4.2.d, would allow the user to communicate in a set of contexts with the support of the application, for example going to the supermarket, to the pharmacy, among others, being those contexts represented in the menu page of Perform Task interface. The Search Word page, Figure 4.2.e, would allow the user to choose from three options, who would help find a word he may want to say but couldn’t because of his diagnosis, being them “White Board” option, “Images” option and “Point Camera” option.

The Aphasia page, Figure 4.2.f, would present the user three options, the first “About Us”, containing a brief description of the product, the “About Aphasia” option, describing aphasia for anyone related to the user who may want to know more about the condition, and the option “Rules of good communication”, that would consist of rules to get a more productive and satisfying communication exchange when communicating with someone diagnosed with Aphasia, helping both the user and the other interlocutor to maintain conversations that both have more chances to understand.

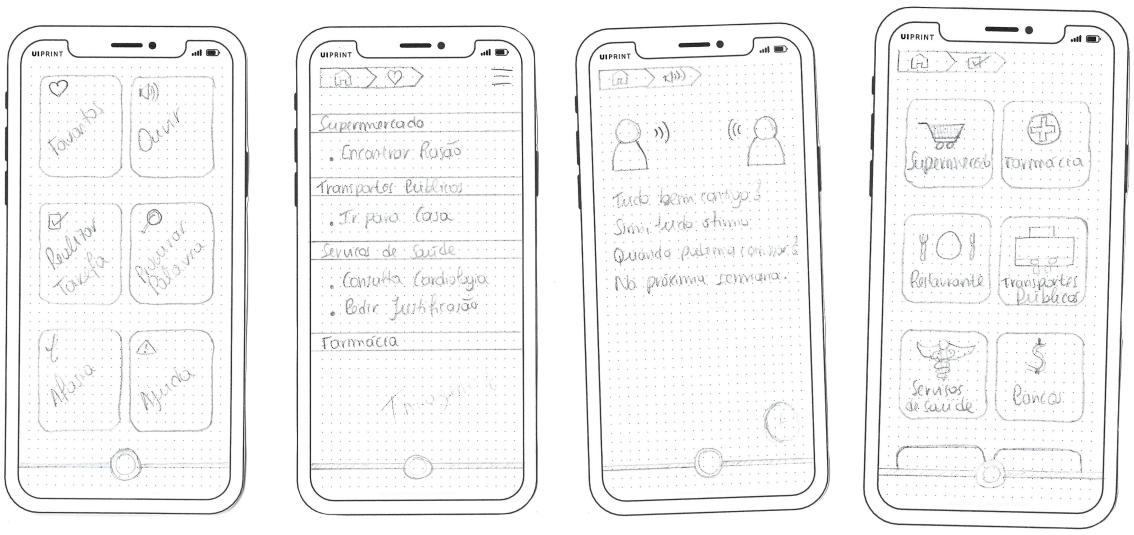
The Help page, Figure 4.2.g, would contain at least three options to be used when the user may feel the urgent need of it, like for example call a predefined emergency contact, send a message to that predefined contact with the exact location of the user or an SOS ready to be displayed if the user felt that he was in an emergency situation. The button for the Help page is strategically located in the lower right corner of the page as it may be more easily within reach of the right thumb of the user.

After an evaluation of this Low-fidelity prototype was performed, some improves were identified, being them:

- In the Menu page, "Listen" would be better replaced by "Speak", as it can be more revealing of its major objective.
- In the Favorites page, and since a multimodal approach is desired, a direct access to what the user saved as favorite, whether that is an image, or a text file or an audio file, divided by the existing contexts in the Perform Task page, would make more sense rather than dividing it by types of file, being that too technical.
- In the Listen page, it would make more sense if the button "Start Listening" was removed and immediately the transcription started as soon as the page was accessed.
- In the Listen page, a floating action button (promoted action for the user) to give direct access to the "Rules of good communication" may be beneficial to the user or other parties in the conversation.
- In the Perform Task page, try to fit at least six options like in the Menu page and showing the user that the page is scrollable would be more suitable.

4.1.3 Third Low-fidelity Prototype

Having in consideration the Evaluation performed on the second Low-fidelity prototype, a third approach of a Low-fidelity prototype was designed and achieved.



(a) Menu page **(b)** Favorites page **(c)** Speak page **(d)** Perform Task page

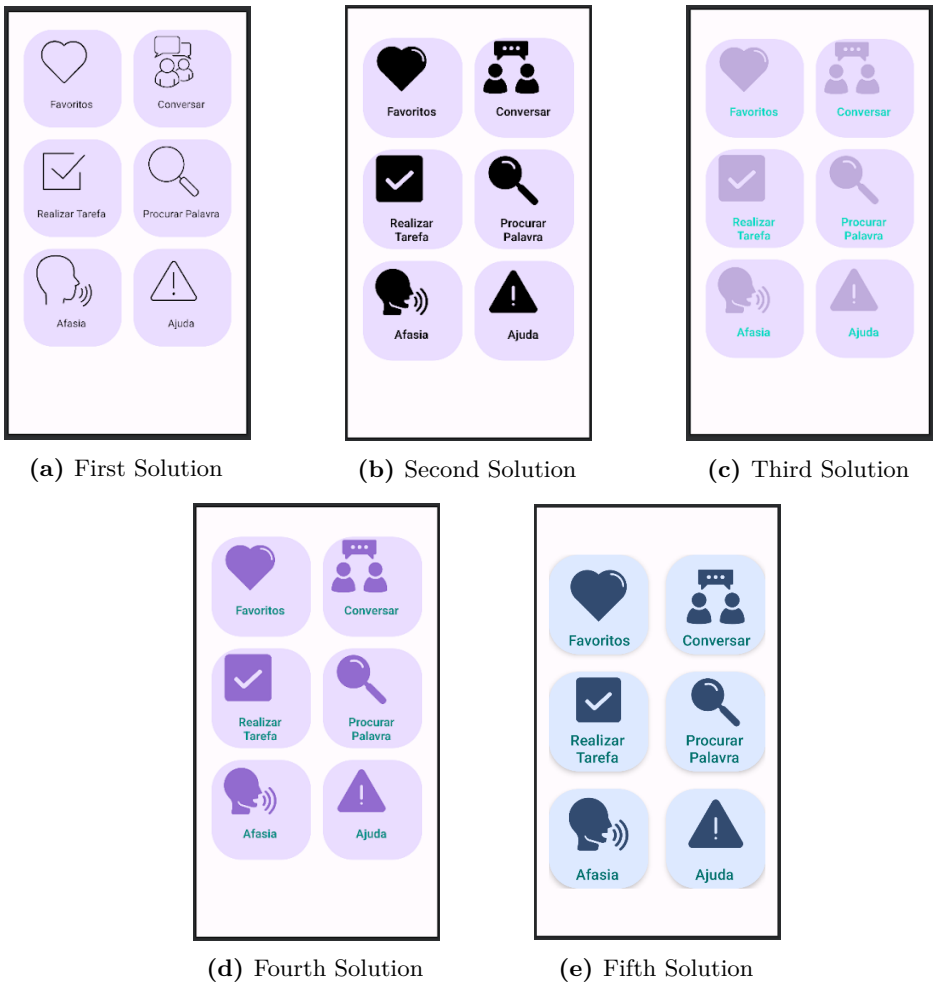
Figure 4.3: Third achieved Low-fidelity prototype, containing the main interfaces of the application, Favorites, Speak, Perform Task, Search Word, Aphasia and Help.

As seen in Figure 4.3.a, in the Menu page, the "Listen" button was replaced by "Speak" button. In the Favorites page, Figure 4.3.b, a direct access to what the user saved as favorite was attempted, whether that is an image, or a text file or an audio file, divided by the existing contexts in the Perform Task page.

In the Speak page, Figure 4.3.c, buttons to start and stop listening were removed with the goal of immediately start to transcript speech. Also in the Speak page, a button was added to allow direct access to the "Rules of good communication". In the Perform Task page, Figure 4.3.d, six options are displayed, showing the user that the page is scrollable.

4.2 Defining Guidelines for Visual Language

Having an acceptable Low-Fidelity Prototype achieved raised the need of defining some visual guidelines regarding the system, once those are not approached in the Low-fidelity outcome.



(a) First Solution **(b)** Second Solution **(c)** Third Solution

(d) Fourth Solution **(e)** Fifth Solution

Figure 4.4: The different five stages of developing visual guidelines for the system.

Having in consideration the Low-Fidelity prototype for the Menu page, present in Figure 4.3.a, and the feedback of a Designer with three years of working experience in application

development, visual guidelines were defined with five developed approaches until a satisfactory solution was achieved. That process is represented in the above Figure 4.4.

In the first solution, Figure 4.4.a, the icons were displayed in buttons as well as the text, in a very simple manner. After the evaluation of the Designer, some observations and suggestions were made, being them:

- Having in consideration "Cognitive Load" (relates to the amount of information that working memory can hold at one time [30]), icons should be filled instead of outlined because they are easier to understand and require less mental effort, and by the same line of thought the font could be bigger in size and thicker.
- The layout is great, white background is always a good choice and also white background is an argument in favor of cognitive load, any colour on white is easier to understand mentally.

Taking into account the feedback given by the Designer relative to the first solution, a second solution was developed, Figure 4.4.b, having the icons become filled instead of outlined and the text font bigger and bold. Again, feedback was given regarding this second solution:

- The filling looks great, but the icons being black makes the solution heavy visually, trying to change the colour of the icon to a matching darker purple of the buttons would be suitable.

Having in consideration the previous given feedback, a third solution was achieved, Figure 4.4.c, changing the colours of the icons to purple and the text to green. The solution was evaluated by the Designer who gave the following feedback:

- Having in consideration the "Web Content Accessibility Guidelines" (four principles: perceivable, operable, understandable, and robust), the icons and background have few contrast. Trying a darker purple would be better.
- Also taking into account the "Web Content Accessibility Guidelines", the text colour is too neon and for that reason trying a darker green would suit best.

Given the previous feedback, a fourth solution was designed, Figure 4.4.d, having the colour of the icons become darker as well as the text colour. Another evaluation took place, having the following outcome:

- The contrast ratio of the icons and the text, given the buttons background, are now under the "Web Content Accessibility Guidelines".
- The buttons should have a dropshadow to give them a 3D effect, which is also in favor of "Cognitive Load".
- Centering the buttons in the page and the icons and text in the buttons would benefit the design.
- Given the context of the application, and by the same logic that was used until now, trying an approach of more sober colours would be suitable.

The last and fifth solution was achieved, Figure 4.4.e, after the previous feedback was taken into account, having an approach of more sober colours being considered. In the buttons, a dropshadow was added as well as all of them were centered and also the icons and text in them. The solution was evaluated by the Designer who gave the following feedback:

- Colours are more sober and the contrast ratio of the icons and the text are under the "Web Content Accessibility Guidelines".
- The dropshadow fits the solution much better.
- Overal, everything is now fitting and suitable regarding the design.

Having the last feedback into account, it is considered that an adequate solution was reached regarding the goals of the present section.

4.3 Conclusions

As a first step in developing a communication multi-platform, an acceptable Low-fidelity prototype of the system was created to validate the paradigm and overall features, having that being achieved through a process of three redefined solutions.

After that, the need to define visual guidelines regarding the system urged and those were defined and refined through a process of five stages of development, leading to a solution that fits the goals of this chapter.

The combination of evaluations and results of the previous sections puts the project in a position to move towards the development of a functional application, which will be detailed in the next chapter.

CHAPTER 5

Iterative Development of Afluentia

In the previous chapter we designed and iteratively refined a low fidelity prototype of the system in order to validate the paradigm and overall features. Also visual guidelines regarding the system were defined. With this achieved, the project is underway to move towards the development of the system itself.

This chapter discusses the used methods to implement the previous defined requirements. The system architecture is presented as well as the technical options that were taken. The current version of the application is depicted, with most of the features mentioned in chapter 3 implemented.

At the end of this chapter an iterative evaluation phase with several stages is described in detail, that will led to the final version of Afluentia.

5.1 Functionalities associated with the Identified Features

As mentioned before, after an acceptable Low-fidelity prototype of the system was achieved as well as visual guidelines regarding the system were defined, as described in chapter 4, the decision to move towards the development of the system itself was made.

From all the available options to develop a native mobile application that allows the iterative development of the solution, already highlighted in chapter 2, Android was chosen.

Figure 5.1: Mobile Operating System Market Share in Portugal as of June 2022. Data collected from the website *StatCounter* [31].

According to StatCounter, which is a web traffic analysis website, and as depicted in Figure 5.1, 73.92% of users of mobile devices in Portugal, as of June 2022, use a mobile device whose operating system is Android [31]. With the aim of reaching as many users as possible,

that was one of the main reasons to develop for Android, in addition to all the advantages inherent to it.

In regards of the programming language to develop software for Android, Java was preferred since there was already a background in Java Android development thus facilitating the process of adaptation and given the spectrum of opportunities, once there is a huge community of Java programmers, and that means it is easier to get answers to critical issues in run time, given that facing technical problems is a common scenario [32].

Android Studio's Integrated Development Environment (IDE) was chosen because has numerous advantages as for example the fact that it comes with a Layout Editor, that is a visual drag and drops editor tool for working with XML files, which helps to create an entirely app layout with ease reducing time of development. Also it comes with Firebase Support with possibility of adding many services such as app analytics, authentication, notification messages as well as with Integrated Cloud that can easily integrate the app with the Google Cloud platform.

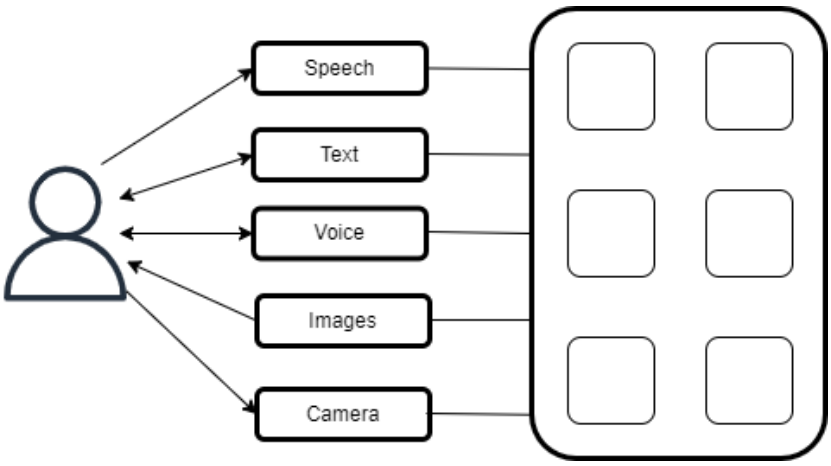


Figure 5.2: Main features of the system identified from the Requirements, being them Speech, Text, Voice, Images and Camera.

When analyzing the Requirements present in chapter 3, it is possible to identify the priority requirements of the system, and associated to such requirements it is possible to distinguish five main groups of features to be part of the system, being them Speech, Text, Voice, Images and Camera, like depicted in Figure 5.2. Associated to such features there are plenty of services that need to be explored and brought to discussion.

Functionalities associated with the already mentioned main features are going to be highlighted. Starting by exploring the Speech feature, it is possible to associate two functionalities in here, **Text to Speech** also called synthetization, a technology that converts text to voice, and **Speech to Text** or transcription, a technology that enables human speech to be converted automatically into text. To the Voice feature it is possible to associate the **Voice Recording** functionality, that will enable to record voice from a microphone, and **Voice Reproduction**, a re-creation of sound waves such as spoken voice. Regarding Images, this feature is easily associated with the **Sketch Recognition** functionality, that will enable

the identification of a Sketch in text form, as well as the **Word to Image** one, enabling the depiction of Images representative of such Words. For the Camera feature, the **Object Detection** functionality, that identifies Objects in a photo taken with the device's Camera can be here associated. The Text feature can be associated with text **Translation**, that will occur if any type of recognition models only recognizes in English and the need to traduce that to Portuguese arises, once the system is for Portuguese speaking aphasic users.

Lastly, one or more **Storage** services will be needed once the user is supposed to have the possibility to access previously saved audio, text and image files, as is possible to extract from the Requirements. Technologies and services associated to the previous identified functionalities were explored, as described in Appendix B, and led to the choices made for the system architecture.

5.2 Overall System Architecture

Having all the technologies been explored, depicted in Appendix B, as well as the previously mentioned prototype and visual guidelines, decisions regarding such technologies were taken and the system was developed. The architecture achieved is based on services, some of which were previously explored.

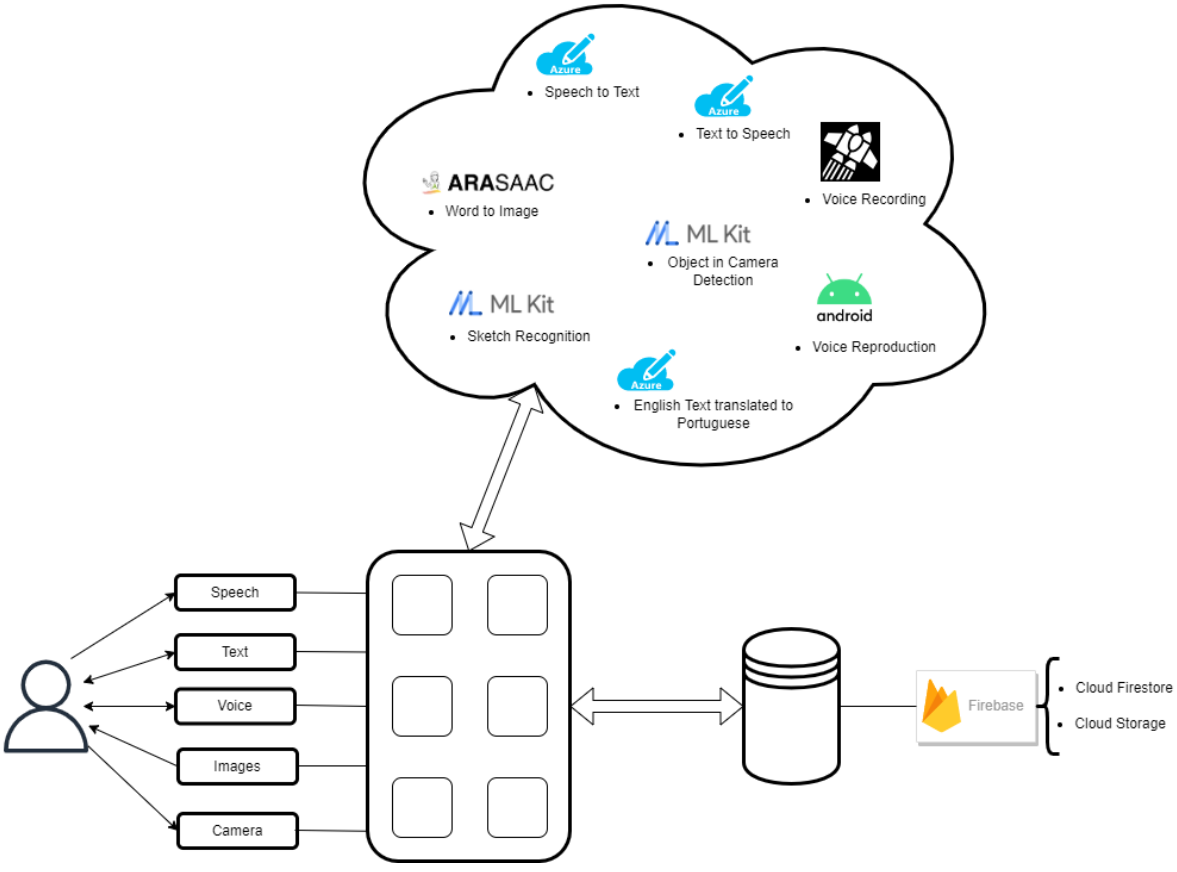


Figure 5.3: Architecture Diagram, based on services, of Afluentia.

Figure 5.3 depicts the architecture of the developed system. As it can be seen, for the **Text to Speech** also called synthetization functionality as well as for the **Speech to Text** or transcription, the chosen technologies were **Azure's Text to Speech API** and **Azure's Speech to Text API** respectively, both part of Azure's Speech Service. This option was made once that, even though at first this service is not free, University of Aveiro has protocols with Microsoft allowing any of its users to access most of Microsoft's Azure services freely and also taking into account the good quality and trustable services they offer. Also, excluding Google Cloud's Text-to-Speech API and Google Cloud's Speech-to-Text API was done because both are Google's paid services for which the University has no protocols with. The Android's TextToSpeech API was also implemented but comparing with the Azure's Text to Speech API implementation, it felt less robust and with less possibility of custumization. Regarding Android's SpeechRecognizer API, this option was immediately excluded once that as it is stated in the documentation, it is not intended to be used for continuous recognition, which is one of the goals in order to implement realtime speech recognition and recognition in audio files.

```
// Replace below with your own subscription key
private static final String speechSubscriptionKey = "[redacted]";
// Replace below with your own service region (e.g., "westus").
private static final String serviceRegion = "westeurope";
```

Figure 5.4: Azure's Speech Service provided keys, namely specified subscription key and region.

```
SpeechConfig speechConfig = SpeechConfig.fromSubscription(speechSubscriptionKey, serviceRegion);
```

Figure 5.5: Azure's Speech Service Configuration.

After registering with University of Aveiro credentials in the Azure portal and after selecting the Speech Service, a pair of subscription keys generated by API Management were provided, present in Figure 5.4. Such key pair allows the creation of an instance of a speech configuration with specified subscription key and region, like depicted in Figure 5.5. That configuration is the way to create a SpeechSynthesizer and a SpeechRecognizer, allowing Text to Speech and Speech to Text respectively.

Regarding the **Voice Recording** functionality, at first, the chosen option was the Android's MediaRecorder API once it is a well documented, free and easy to use technology. But MediaRecorder doesn't record audio in WAV format, which became a disadvantage once that to perform transcription from previous saved audio files, Azure's Speech to Text default audio streaming format is WAV (16 kHz or 8 kHz, 16-bit, and mono PCM). After excluding the MediaRecorder API, Android's AudioRecord API was considered, but the fact that it doesn't provide a direct way to write to a file, which is super important to save audio files into storage, didn't make it the best solution either. Since none of the explored solutions before development was an ideal solution, a new technology that would record audio in WAV format and write that recorded audio directly into a file was needed.

After a search for such characteristics, an **Android Wav Recorder library** hosted in JitPack was found, providing the possibility to record WAV/PCM files and the Recorder instance constructor is also composed by a file saving path, allowing a direct way to write to a file.

As far as **Voice Reproduction** is concerned, the chosen technology was the **Android's MediaPlayer API** because like the MediaRecorder API, is well documented, free and easy to use. The difference is that MediaPlayer supports WAV files and because of that, remained the used technology in this project to reproduce audio files.

Regarding **Sketch Recognition**, before development there was only one technology explored, which was **ML Kit's Digital Ink Recognition API** and that remained as the used solution. Very well documented technology and with plenty of sample codes, the AutoDraw model is the used one to detect sketches made by the user, which pairs machine learning with drawings from talented artists to enable AutoDraw's suggestion tool to guess what the user draw.

When it comes to **Word to Image** functionality, and once this is a system for people with fluent types of aphasia, a simple word to image conversion with average images wouldn't be enough and that's why the **ARASAAC REST API for pictograms**, already explored, remained as the used solution once that when compared to other pictographic systems, the ARASAAC database appears to be one of the most appropriate augmentative alternative communication systems [33].

To perform **Object Detection** with the phone camera and considering the previous mentioned options, the used one was **ML Kit's Object Detection and Tracking API**. The two technologies were experimented and the only reason the used one was the ML Kit's one was due to the ease of changing from live detection, stream mode, to detection in static images, single image mode.

Regarding the **Translator** service, and looking at the previous considered solutions, once again Google Cloud's Translation API is a paid service for which the University has no protocols with and because of that is not an option for this project. With that being said, the used option is the **Azure's Text Translation REST API** because first of all University of Aveiro has protocols with Microsoft allowing any of its users to access most of Microsoft's Azure services freely, being the Translator one of them, and again because of the good quality and trustable services they offer.

```
private static String key = "[redacted]";
private static String location = "westeurope";

HttpUrl url = new HttpUrl.Builder()
        .scheme("https")
        .host("api.cognitive.microsofttranslator.com")
        .addPathSegment("/translate")
        .addQueryParameter( name: "api-version",  value: "3.0")
        .addQueryParameter( name: "to",  value: "pt")
        .build();

// Instantiates the OkHttpClient.
OkHttpClient client = new OkHttpClient();

// This function performs a POST request.
public String Post() throws IOException {
    MediaType mediaType = MediaType.parse("application/json");
    RequestBody body = RequestBody.create(mediaType,
            content: "[{\"Text\": \""+toTranslate+"\"}]");
    Request request = new Request.Builder().url(url).post(body)
            .addHeader( name: "Ocp-Apim-Subscription-Key", key)
            .addHeader( name: "Ocp-Apim-Subscription-Region", location)
            .addHeader( name: "Content-type",  value: "application/json")
            .build();
    Response response = client.newCall(request).execute();
    return response.body().string();
}
```

Figure 5.6: Azure's Text Translation REST API instantiation and function that performs a POST request.

The request Post() method sends a POST request to the specified HttpUrl. To connect to the Translator service via the REST API, the headers present in Figure 5.6 need to be included with each request. *Ocp-Apim-Subscription-Key* corresponds to the Translator service key from the Azure portal, *Ocp-Apim-Subscription-Region* to the region of the multi-service resource and *Content-type* to the payload content type, which can only be *application/json* or *charset=UTF-8*.

Finally, and as far as **Storage** is concerned, the options used are **Firebase Storage Services**. It was chosen due to the fact that there was already some background dealing with such technologies and also for having a quick and easy integration and configuration, because Android Studio comes with Firebase Support, as well as concise documentation.

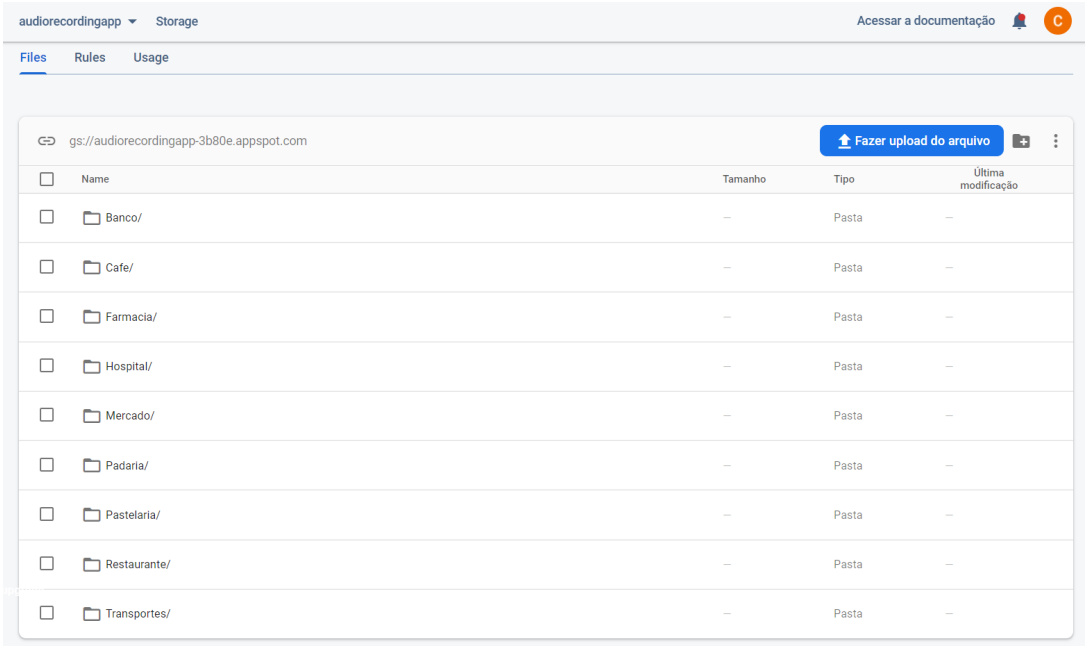


Figure 5.7: Structure of Firebase Cloud Storage project, containing all the context folders that store text files and audio files the user saved.

Cloud Storage, as seen in Figure 5.7, is used in order to store the media files, that in this case are audio and text files from the transcriptions and synthetizations, that the user chose to save and related to the specific contexts.

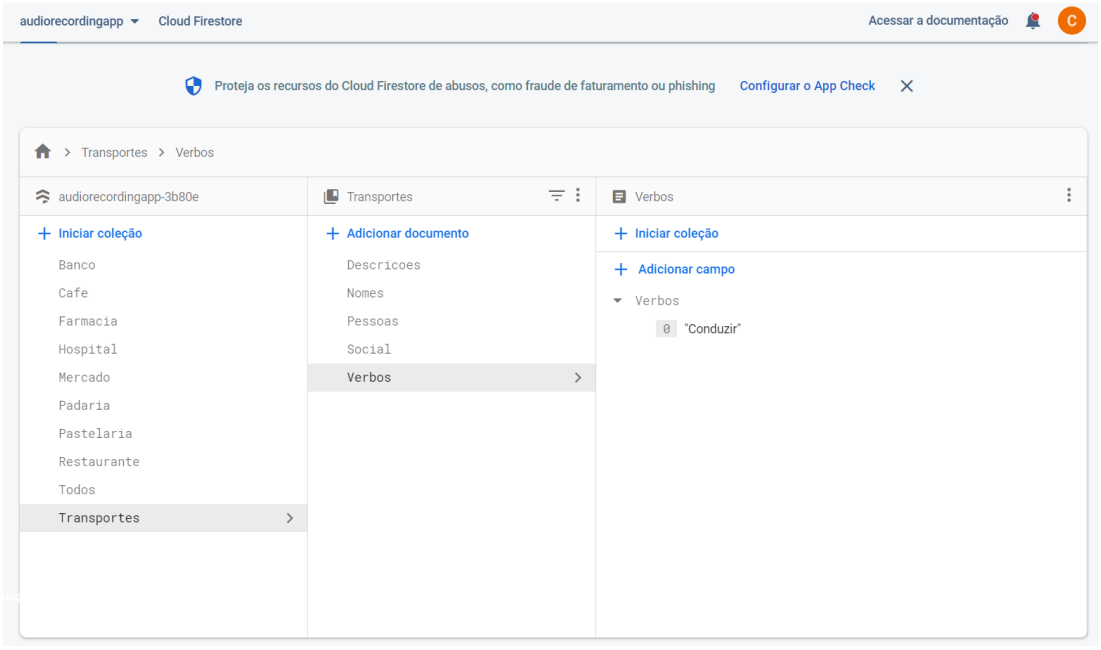


Figure 5.8: Structure of Firebase Cloud Firestore project, containing all the context folders used to to store the words present in the Words page interface.

Cloud Firestore is used to store the words present in Words page, also associated with the given contexts and depicted in Figure 5.8.

5.3 Iterative Development and Evaluation

A first version of the system was developed taking into account priority requirements from the perspective of the project. Looking at the requirements present in chapter 3, the goal was to develop the requirements defined as priority 1 and considered essential. Associated to such requirements, was the development of the Speak page, of the Perform Task page and of the Search Word page as well as all the pages that derive from them.

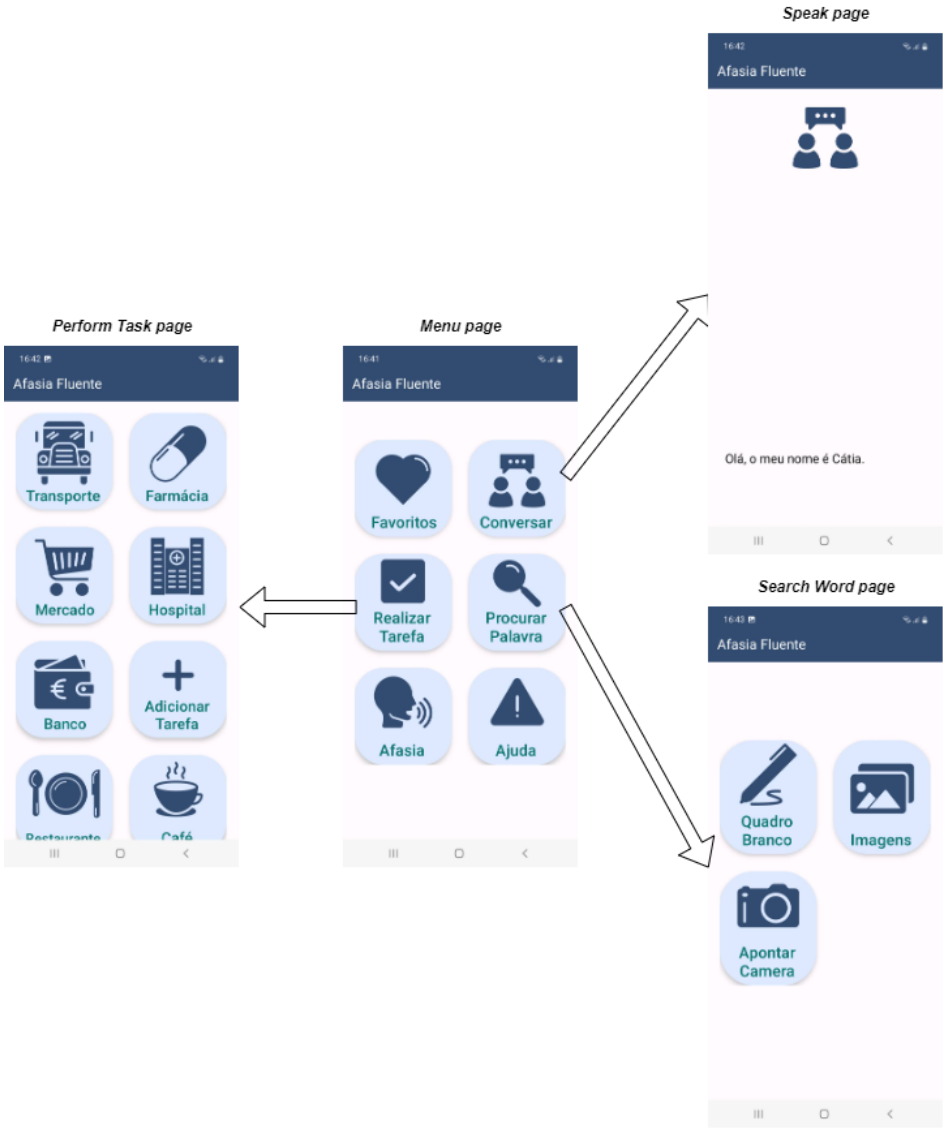


Figure 5.9: First version of the main developed interfaces, that derive from the Menu page, being them the Speak page, Perform Task page and Search Word page.

In the above Figure 5.9, an user interface representation of the main developed interfaces that derive from the Menu page is depicted.

The fulfillment of such requirements was accomplished not only taking into account the already defined low-fidelity prototype and visual guidelines, but also with a continuous evaluation of the solution throughout all the process of the development with a constant

iterative refinement, carried out with a focus group composed by a Speech and Language Therapist (SLT) and a Human Computer Interaction (HCI) Expert, and considering principals of usability aligned with Nielson's heuristics, also already mentioned in this document in chapter 2.

5.3.1 First Evaluation

After a first version of the system was achieved (a complete depiction of all screens is provided in Appendix C), a guide with not only questions but predefined tasks relative to the application was prepared in order to start an evaluation phase with structured focus groups. Such guide is present in Appendix D. Looking at it, it is possible to notice that for every identified main interface, Speak page, Perform Task page and Search Word page as well as for the interfaces deriving from it, such as White Board, Images and Point Camera, it is made a contextualization once not all members of the focus groups may be familiar with the app and in what it consists of.

The decision to start the evaluation phase with experts in this subject and not with the end users for whom this application is intended, was made taking into to account that the target users suffer from a condition, which in this case is some type of fluent aphasia and they may be weakened, leading to possible application problems be interpreted by them as their own problems. Overall respect for the user was privileged and prioritized, and test and evaluate the app with end users will be performed when major issues have been solved and when a level of satisfaction is consensual by all those involved in this phase.

Evaluation

Every task that is presented in Appendix D was intended to be evaluated by the focus group, being asked to give an evaluation from 1 to 5, where 1 is too difficult and 5 very easy, and also by me, using observational techniques and considering parameters like what the participant would say and do, noting places where difficulties were encountered, analysing how the task was accomplished regarding task duration, frequency, complexity, surrounding conditions and any other factors involved in or required to perform the given tasks, all of that without interfering in the participant's actions.

The first evaluation that took place was carried out with a Speech and Language Therapist, being 33 years old and having 6 years of working experience, inclusively with people with aphasia, that is part of the APH ALARM project, in which this dissertation inserts in. The Speech and Language Therapist was never in contact with the app until that moment and because of that, a contextualization regarding the app and the dissertation itself was made. After that, and in order to understand if the flow of the application was adequate and perceivable for the target users and also to familiarize the SLT with the system, the opportunity to explore the application without any tasks was given.

Results and Discussion

Since no problems were identified and no suggestions were made while the SLT was exploring the system, the guide with the question and predefined tasks took place.

	Speech Therapist	Me
Conversar - Tarefa 1	4	4
Realizar Tarefa - Tarefa 1	5	4
Realizar Tarefa - Tarefa 2	2	2
Realizar Tarefa - Tarefa 3	2	2
Realizar Tarefa - Questão 4	*"It may vary depending on the type of aphasia."*	-
Realizar Tarefa - Tarefa 5	5	5
Procurar Palavra Quadro Branco - Tarefa 1	5	5
Procurar Palavra Imagens - Tarefa 1	2	2
Procurar Palavra Apontar Câmera - Tarefa 1	4	4

Table 5.1: First Evaluation performed after Development, carried out with a Speech and Language Therapist and evaluated with a scale from 1 to 5, where 1 is too difficult and 5 very easy.

In Table 5.1 it is possible to see the results of the evaluation present in Appendix D. After every task was accomplished, some suggestions and considerations were made by the SLT relative to some of the tasks, and notes of such suggestions were taken. Relative to Task 1 of the interface Speak (**Conversar - Tarefa 1**), for which the given evaluation by the SLT was 4, meaning that was easy to perform, and also 4 by observation, meaning that the SLT performed the task with no apparent difficulties, the following considerations were made:

- *"The recognition works very well."*
- *"Adjust the font size of the transcription to bigger, or give that possibility to the user."*
- *"Consider a greater spacing between words. With the increase of the font size it may not be necessary."*
- *"Different colors for different recognized voices would allow the user to identify all the participants easily."*
- *"Placing the speech of the user (person with aphasia) aligned to the right and the speech of the other participants aligned to the left would also allow the user to better follow the conversation."*
- *"In the end of the conversation, a summary with all the main addressed topics."*
- *"Fitzgeral Key words underlined and with colors. If the user selects them, images representative of those words appear"*

Relative to Task 1 of the interface Perform Task (**Realizar Tarefa - Tarefa 1**), the given evaluation by the Speech and Language Therapist was 5 and thus considered very easy to perform and 4 by observation, not noticing any major strugglings performing the given task. To this task, the following consideration was made by the SLT:

- *"Increase the font size of the existing list of tasks."*

In regard of Task 2 of the interface Perform Task (**Realizar Tarefa - Tarefa 2**) the evaluation given by the SLT was 2, which means that difficulties were found when performing the task, and by observation the same evaluation was given once that it was perceivable the struggle the SLT made to accomplish what was intended. To this task, the following considerations were made:

- *"The order and meaning of the buttons is not very suggestive, some can be miss perceived with others and cause confusion to the user. Try a new approach to the page flow."*
- *"Placing a small text at the top, explaining the purpose of this interface would help the user to better understand it."*
- *"Place icons on buttons."*

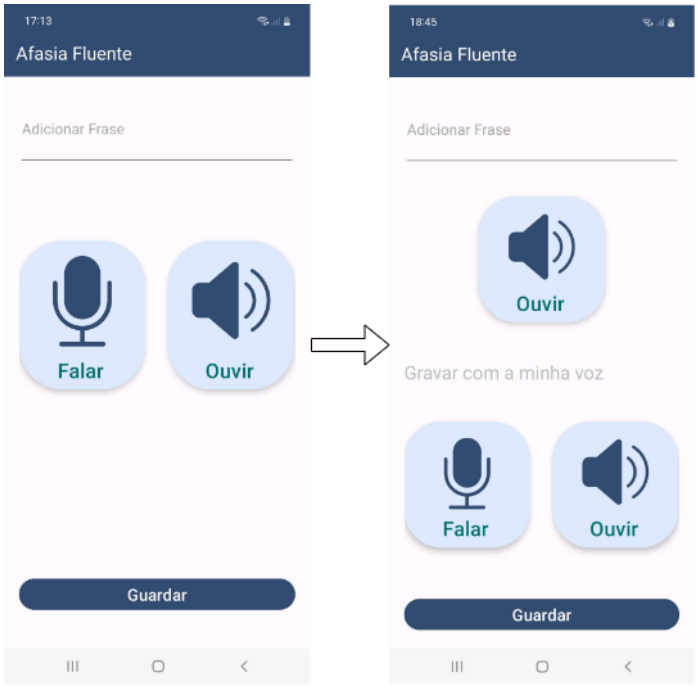


Figure 5.10: Redefined Layout of Add Task page after the first evaluation, separating the Write and Speak interfaces.

After the first suggestion was considered, a new page flow attempt was performed and the redefined layout is present in Figure 5.10, with the part where the user is supposed to write being separated from the part where the user is supposed to speak, attempting not to cause miss perceptions to the user.

Relative to Task 3 of the interface Perform Task (**Realizar Tarefa - Tarefa 3**) the evaluation and considerations that were made are the same from Task 2 of the interface Perform Task (Realizar Tarefa - Tarefa 2) once their interface is the same. Relative to Task 5 of the interface Perform Task (**Realizar Tarefa - Tarefa 5**) the given evaluation by the SLT and by observation was 5, considering that it was very easy to perform the task. Nonetheless, the following considerations were made:

- *"When the task is opened, increase the the font size."*

- *"Put icons on the button."*

Talking about Task 1 of the interface Perform Task and sub interface White Board (**Procurar Palavra, Quadro Branco - Tarefa 1**), the evaluation given by the Speech and Language Therapist was 5 and by observation too, with no difficulties in performing the task being noticed. Despite that, the following considerations were made by the SLT:

- *"Place icons on buttons to make it easier to understand their meaning."*
- *"Change word "text" to "write" or to "word"."*
- *"Try to come up with a new word for the button 'recognise'."*

Talking about Task 1 of the interface Perform Task and sub interface Images (**Procurar Palavra, Imagens - Tarefa 1**), the evaluation given by the SLT was 2 and the same evaluation bye observation was given, once that some difficulties performing the task were noticed. The following considerations were made by the Speech and Language Therapist:

- *"Too many menus, possibly remove the menu where words are filtered by Fitzgerald Key and just put that as a filter in the list of all words."*
- *"Place explanatory text at the beginning so that the purpose of this interface is more perceptible to the user."*

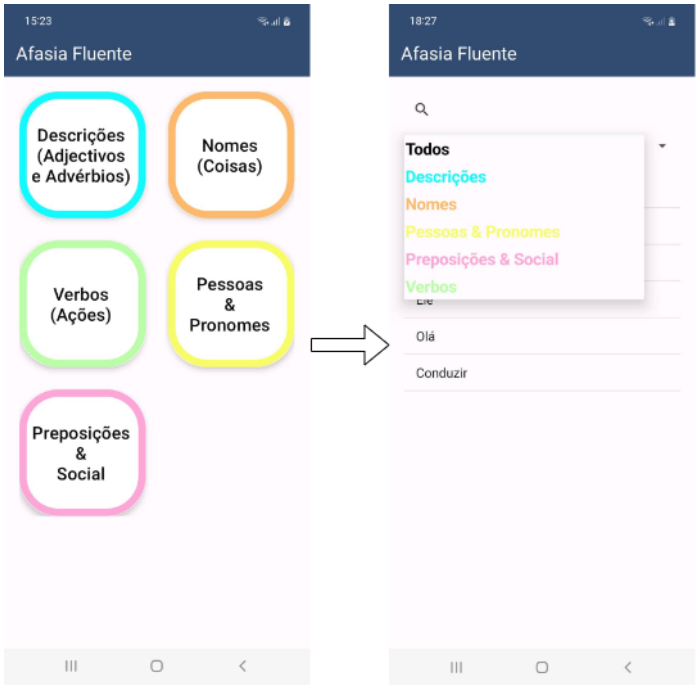


Figure 5.11: Removal of Fitzgerald Key page and redefined layout of Words page after the first evaluation, reducing the number of menus and decisions a user has to make.

After the first suggestion being taken in consideration, a new attempt of the page flow was accomplished and the redefined layout is present in Figure 5.11, with the removal of the Fitzgerald Key page and consequent substitution by a filter present in the Words page, which decreases the number of menus and immediate decisions the user has to face.

Talking about Task 1 of the interface Perform Task and sub interface Point Camera (**Procurar Palavra, Apontar Câmera - Tarefa 1**), the SLT evaluated this task with 4 and the same evaluation was given by observation once that it seemed to be pretty easy for the SLT to do what was intended. The following consideration was made:

- *"Instead of a dynamic image, in which several objects can be identified, it would be more beneficial and make more sense to have a static image since the objective of this interface is to find only one word and/or object and not several."*

The previous suggestion was taken into account and the development of object detection in a static image was performed, allowing the user to take a photo of the intended object instead of pointing with the camera to it.

5.3.2 Second Evaluation

As an outcome of the previous evaluation, requirements were refined taking into account all the suggestions and considerations made. The redefinition of the requirements lead to improvements in the application, and as a consequence of that, the need to perform another evaluation urged, in order to validate not only what was already done but also the performed changes and improvements.

Once again, the evaluation followed exactly the same guide as the previous evaluation, as well as the same evaluation techniques, that is, all the tasks presented in Appendix D were intended to be evaluated by the focus group, evaluating from 1 to 5, where 1 it is very difficult and 5 very easy, and also by me, once again using the same Observational Techniques already mentioned, all without interfering with the participant's actions.

The evaluation was carried out with another Speech and Language Therapist and teacher at University of Aveiro, being 42 years old, with 20 years of working experience and few experience of working with people with aphasia. Once the teacher was already aware of this dissertation as well as aware of the realization of the app, having never had contact with it, no contextualization was made and the opportunity to explore the application without any tasks was firstly given, once again, in order to understand if the flow of the application was adequate and perceivable for the target users as well as to familiarize the teacher with the system.

Results and Discussion

By exploring the app, some first considerations and suggestions were made:

- *"An initial tutorial to be displayed the first few times a user would use the app, explaining what the purpose of each menu was, could help the user to better understanding of the app and know all the available options."*
- *"The "Favoritos" icon at the Menu page should not be the first icon once that at the beginning is empty.'*
- *"Good choice of the app colors. Often, AAC apps offer some childish colors, leading some users to feel less comfortable to use the systems."*

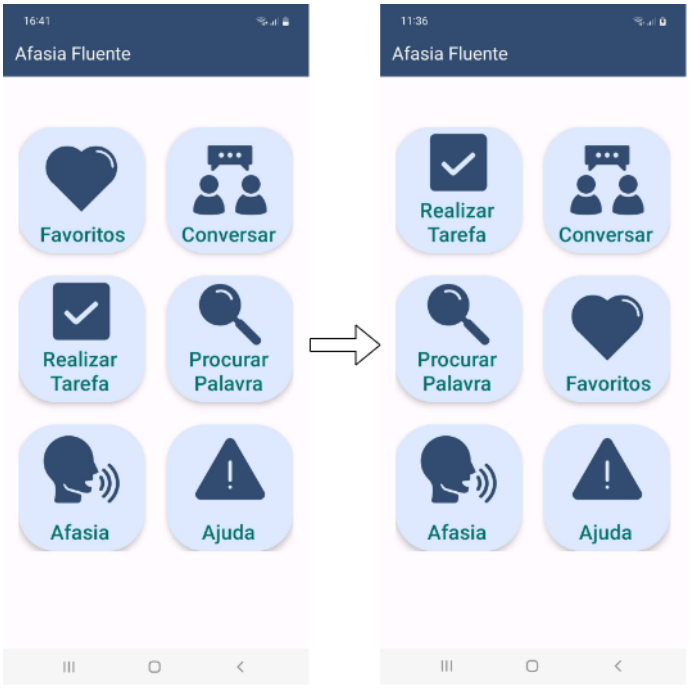


Figure 5.12: Redefined layout of Menu page after first impressions of the second evaluation, reallocating the "Favoritos" Button.

Having the second suggestion been taken in consideration, a new attempt of the Menu page flow was accomplished and the redefined layout is present in Figure 5.12. Once the teacher finished exploring the app, and since no problems were identified, the guide present in Appendix D took place.

	Speech Therapist	**Me**
Conversar - Tarefa 1	4	4
Realizar Tarefa - Tarefa 1	5	5
Realizar Tarefa - Tarefa 2	2	2
Realizar Tarefa - Tarefa 3	2	2
Realizar Tarefa - Questão 4	*"Depends on the type of aphasia the user has."*	-
Realizar Tarefa - Tarefa 5	5	5
Procurar Palavra Quadro Branco - Tarefa 1	4	5
Procurar Palavra Imagens - Tarefa 1	5	4
Procurar Palavra Apontar Câmera - Tarefa 1	4	4

Table 5.2: Second Evaluation performed after Development, carried out with a Speech and Language Therapist and teacher at University of Aveiro and evaluated with a scale from 1 to 5, where 1 is too difficult and 5 very easy.

In Table 5.2 it is possible to see the results of the previously mentioned evaluation (a full account of the outcome are provided in Appendix D). Relative to Task 1 of the interface Speak (**Conversar - Tarefa 1**), the given evaluation by the SLT and observation was 4, that is, considered easy to perform. The following considerations were made by the Speech and Language Therapist:

- *"Notifications and/or alerts for when a conversation participant is talking too fast, which may be difficult for a person with aphasia to understand and/or read."*
- *"Add scrollbar to show that it is possible to move back."*
- *"When the user moves backwards, do not move to the beginning when a voice is recognized again. Only by user's will."*
- *"Add explanatory text at the beginning of the interface."*

Relative to Task 1 of the interface Perform Task (**Realizar Tarefa - Tarefa 1**) no suggestions or considerations were made and the task was considered very easy to perform by not only the Speech and Language Therapist but also by observation. Relative to Task 2 of the interface Perform Task (**Realizar Tarefa - Tarefa 2**), and as in the previous evaluation, the task was considered by the SLT difficult to perform and observation concluded the same. The following considerations were made:

- *"Before entering the page, the button to access the same "Add Task" is very technical. Try a better approach, possibly "Add another" or "Add more" or even "Add activity", all of which are more illustrative of an everyday action."*
- *"As for the page itself, there is too much information present on it, which makes it not very intuitive and not very perceptible in relation to what to do. The best thing would be to separate the two interfaces "Write" and "Talk" on different pages."*

After the second suggestion was taken in consideration, a new attempt of the page flow was accomplished and the redefined layout is present in Figure 5.13, separating the Add Task page in two different pages, the Speak and Write pages.

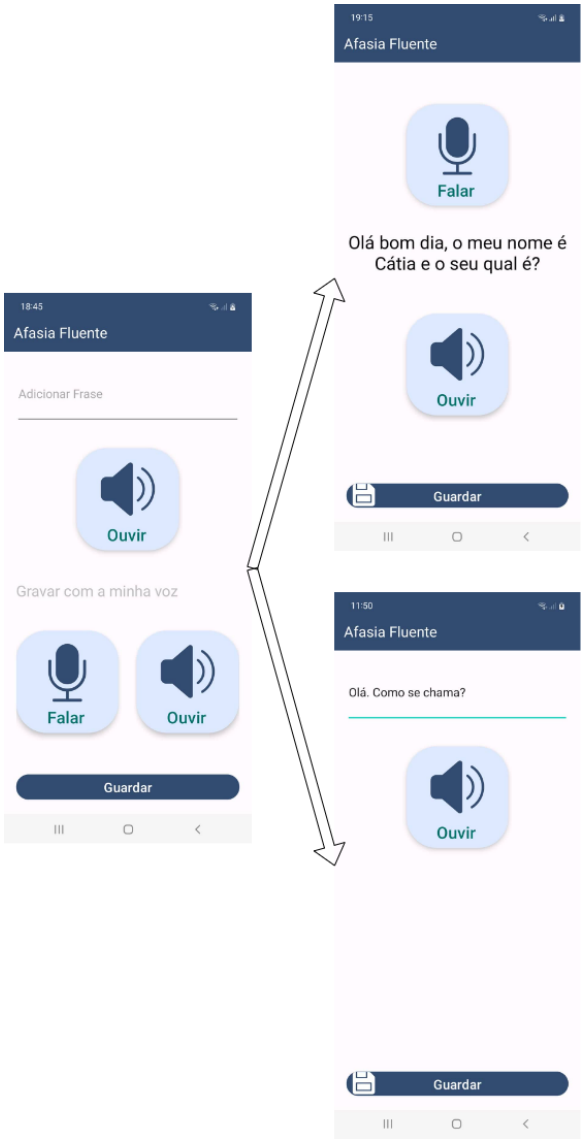


Figure 5.13: Removal of single Add Task page and separation into Speak and Write interfaces after the second evaluation.

Relative to Task 3 of the interface Perform Task (**Realizar Tarefa - Tarefa 3**) the considerations that were made are the same from Task 2 of the interface Perform Task (Realizar Tarefa - Tarefa 2) once their interface is the same. Relative to Task 5 of the interface Perform Task (**Realizar Tarefa - Tarefa 5**) no suggestions or considerations were made and the given evaluation was 5 by both Speech and Language Therapist and observation, having no difficulties been found performing the intended. Talking about Task 1 of the interface Perform Task and sub interface White Board (**Procurar Palavra, Quadro Branco - Tarefa 1**), the evaluation given by the SLT was 4 and 5 by observation, once that no difficulties performing the task were noticed. In regard of this task, the following considerations were made:

- *"Place the "Reconhecer" and "Limpar" buttons at the bottom of the interface, along with the "Ouvir" button."*
- *"Try to come up with a new word to "Reconhecer", for example, "O que é isto?"."*

Talking about Task 1 of the interface Perform Task and sub interface Images (**Procurar Palavra, Imagens - Tarefa 1**), no suggestions or considerations were made once it was acknowledge to be a very straightforward and easy to accomplish task, with a given evaluation of 4 by the SLT and 5 by observation. Relative to Task 1 of the interface Perform Task and sub interface Point Camera (**Procurar Palavra, Apontar Câmera - Tarefa 1**), it was considered an easy to accomplish task with a given envaluation of 4 by both parties. The following considerations were made:

- *"After the photo is taken, and while the image doesn't load, add a loading icon to the interface."*
- *"After the image loads, put the "OK" button in green."*

Once the previous two evaluations were successfully concluded, all the suggestions and considerations were taken into account and requirements were refined, leading to some readjustments whether regarding technical options or layout ones, the final user interfaces of the application was achieved (a complete depiction of all screens is provided in Appendix E).

5.4 Conclusions

In the beginning of this chapter, all conditions were reunited to start the implementation and development of the Afluentia application. For that, it was necessary to define an architecture based in services and representative of all that had been discussed until that moment, scalable and modifiable enough so that in the future the possibility of continuous development and improvement exists.

From the requirements present in Chapter 3, it was possible to identify the priority requirements of the system, and associated to such requirements it was possible to distinguish five main groups of features to be part of the system, being them Speech, Text, Voice, Images and Camera, which led to plenty of functionalities and further services and technologies to be firstly explored and brought to discussion. Having the technologies been explored, as well as the previously mentioned prototype and visual guidelines defined, decisions regarding such technologies were taken and explained and the system started to be developed.

After the architecture implementation was explained and taking into account the continuous evaluation of the solution throughout all the process of the development with a constant iterative refinement, carried out with a focus group composed by a Speech and Language Therapist (SLT) and a Human Computer Interaction (HCI) Expert, we arrived at a functional version of Afluentia. Finally, this version went through an evaluation phase carried out with focus groups composed by experts in this matter, which led to the current version of Afluentia, present in Appendix E.

CHAPTER 6

Conclusions

In this chapter, the conclusions of this dissertation are presented, where a general analysis of the work carried out is performed, as well as a look into future work and other developments that can further enrich the communication system are taken.

6.1 Outcomes of the developed work

In the beginning of this document it was stated that identifying the communication challenges that people with aphasia face in their routine in order to characterize the motivations and specificities of the patient with aphasia was an objective, as well as to propose solutions that can assist these patients in their day-to-day communication needs. To accomplish such objectives, some goals had to be met.

Firstly, to **gather knowledge regarding aphasia and different existing assistive communication solutions** research was performed. Concerning aphasia, the research performed focused not only on literature but also on feedback and conversations with experts. After revising a wide amount of existing solutions that might serve the intended goals and that were reported as having been used by people with aphasia, we concluded that several challenges and contexts were still not addressed, motivating the proposal of a new approach for people with aphasia.

As a result of the done research and gathered knowledge and information, and **making use of a User-Centered Design methodology**, aphasia patients were modeled into **Personas** and how the system may help them tackle some challenges in particular contexts into **Scenarios**, all validated by Speech and Language Therapists (SLT) and by a Human Computer Interaction (HCI) Expert. One really important step of this work was the development of the Personas, that to our knowledge, are the first to be proposed. They are an important contribution to improve our understanding of aphasia and can serve as grounds for other works.

After achieving a solid family of Personas for Fluent Aphasia as well as the respective Scenarios, and still having in mind the User-Centered Design methodology, objects, actions

and contexts that best define the needs that the product must satisfy were extracted from the Scenarios in order to define the system **Requirements**.

What follows is the prototyping phase. To this end, we decided to start by developing a low-fidelity, hand-drawn prototype, given the constant and rapid changes that this type of prototypes face, since they are a very primary version of what is intended and knowing that their objective is to be fast and efficient. This prototyping phase went through three phases of evaluation carried out by a Focus Group composed of a Speech Therapist and an Expert in Human-Computer Interaction, with the objective of defining in a very general way how the different functionalities would be organized. After reaching an acceptable Low Fidelity Prototype, the need to define some visual guidelines arose and this process had five phases, all of which were evaluated by a Designer with three years of experience in mobile application development.

Taking into account the priority requirements, it was possible to identify five modalities associated with such requirements as part of the system, namely Speech, Text, Voice, Images and Camera. Associated with these modalities are features such as Text to Speech, Speech to Text, among others, for which we use services such as APIs from Azure or Google, which can therefore be considered an architecture based on Services. Finally, the need for storage also arose, as we made use of Cloud Firestore and Cloud Storage.

That being said, the conditions were met to begin the development of the system itself and it was possible to **propose technology-based solutions** for assisted/alternative communication according to the defined Scenarios and extracted Requirements.

Carrying out an **iterative design, development and evaluation of a proof-of-concept solution for communication aspects** was then possible, with all the Requirements defined as priority level 1 being implemented. The proposed solution went through an evaluation phase carried out with focus groups composed by experts in this matter, namely Speech and Language Therapists with a great amount of years of experience, which led to the current version of Afluentia with a good feedback. Having in consideration the goals for this dissertation, the work carried out allowed accomplishing them with success.

6.2 Future Work

As previously said, the main goal of this project, for which this is the first stage, is to achieve an efficient and user centered solution for assisted and alternative communication of people with aphasia. Despite the fact that it has been only tested and evaluated with experts in controlled contexts, it can be concluded that the presented proof-of-concept served its overall purpose, showing the potential and viability of such system. In this context, other developments that can further enrich the communication system will follow, such as:

- **User management**
  Develop user management, allowing users to register in the app, log in and log out. This can be possible done through Firebase Authentication, which provides backend services, easy-to-use SDKs, and ready-made UI libraries to authenticate users to applications. It

supports authentication using passwords, phone numbers, popular federated identity providers like Google, Facebook and Twitter, and more. The possibility to customized information and features to each user would be given.

- **Implement explanatory tutorial to the app or explanatory texts to the interfaces**
  An initial tutorial to be displayed the first few times a user would use the app, explaining what the purpose of each interface was, could help the user to better understanding of the app and know all the available options or explanatory texts present at each interface. All validated with an expert in this matter.
- **Notifications for the Transcription of Speech in real time**
  In the 'Conversar' interface when speech is being transcribed, issue notifications or alerts when the speech rate is too fast (which can greatly diminish a person with aphasia's ability to comprehend or read), notifying the other party that is speaking too fast and alerting to slow down. Firebase Cloud Messaging (FCM) provides a reliable and battery-efficient connection between server and devices that allows to deliver and receive messages and notifications on iOS, Android, and the web at no cost.
- **Implement a summarization method for a conversation from the 'Conversar' interface**
  One of the unimplemented requirements refers to allow the user to get a summary of the conversation he just had at the 'Conversar' interface, overcoming any kind of comprehension or reading difficulties that may have occurred during the conversation, and was not overcome at the time. There is a Python library called sumy for extracting summary from HTML pages or plain texts, which tries to find the most significant sentences in the plain text and compose it into the shortened text.
- **Identification of Fitzgerald's Key words present in the summarized conversation**
  Identification of Fitzgerald's Key words present in the summarized conversation, highlighting those and if the user selects them, images representative of such words appear. The images representative of such words can be retrieved from the REST API for Arasaac pictograms, as it is done in the Image for Word interface.
- **Different voices recognition at the 'Conversar' interface**
  The possibility to identify different voices participating in the conversation would allow the accomplishment of many suggestions made during the evaluation phase, for example placing the speech of the user (person with aphasia) aligned to the right and the speech of the other participants aligned to the left as well as different text colors for the different recognized voices.
- **Clickable timeline**
  Develop a clickable timeline for every interface, allowing the user to situate himself in the app as well as to go to any of the interfaces present in the timeline in a direct way. This can be developed with Android breadcrumbs, that will show a timeline of navigation events.

- **Possibility to add new Context relative to the Perform Task interface**
  Develop the possibility for the user to add a new Context in the Perform Task interface, customizing this interface according to its preferences. This can be achieved by developing user authentication for the app, already mentioned, and associate this type of data with the specific user.
- **Develop the 'Favoritos' interface**
  Develop the possibility for the user to save images, texts and audios as favorites with given access at any time. This can be achieved by developing user authentication for the app, already mentioned, and associate this type of data with the specific user.
- **Develop the 'Afasia' interface**
  Develop the 'Afasia' interface with accessible information like rules of good communication as well as information relative to aphasia. Once again, such information should be verified from an expert in this matter.
- **Develop the 'Ajuda' interface**
  Develop the 'Ajuda' interface with the possibility to ask for immediate help to the emergency contact or display a help message to whoever the user is with.

These are only a few examples of features that could be added in the future to improve Afluentia and make it more exciting and with the aim of embracing more users. It also would be interesting to conduct tests with fluent aphasia diagnosed individual's, our end users, once that we trust that the current version has a set of features capable of doing so and which would further approximate our proof of concept to an application that could be used on a large scale.

APPENDIX A

SUS and PSSUQ Questionnaires

Originally created by John Brooke, the The System Usability Scale (SUS) was described by himself as a "quick and dirty" measurement tool, and it consists of a 10 item questionnaire with five response options, from Strongly disagree to Strongly agree:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system [28].

The Post-Study System Usability Questionnaire (PSSUQ) consists of a 16 item questionnaire (18 questions in Version 1; 19 questions in Version 2; 16 questions in Version 3) with seven response options, from Strongly disagree to Strongly agree, which compared to SUS allows more specific responses to each question:

1. Overall, I am satisfied with how easy it is to use this system.
2. It was simple to use this system.
3. I was able to complete the tasks and scenarios quickly using this system.
4. I felt comfortable using this system.
5. It was easy to learn to use this system.
6. I believe I could become productive quickly using this system.
7. The system gave error messages that clearly told me how to fix problems.
8. Whenever I made a mistake using the system, I could recover easily and quickly.
9. The information (such as online help, on-screen messages, and other documentation) provided with this system was clear.

10. It was easy to find the information I needed.
11. The information was effective in helping me complete the tasks and scenarios.
12. The organization of information on the system screens was clear.
13. The interface of this system was pleasant.
14. I liked using the interface of this system.
15. This system has all the functions and capabilities I expect it to have.
16. Overall, I am satisfied with this system [29].

APPENDIX B

Explored Technologies for the identified Functionalities

B.1 Speech

Regarding the **Text to Speech** functionality the following technologies were explored and considered:

- **Android's TextToSpeech API**
  Synthesizes speech from text for immediate playback or to create a sound file, and can be customized to the Portuguese language.
- **Google Cloud's Text-to-Speech API**
  Google Cloud Text-to-Speech enables developers to synthesize natural-sounding speech with 100+ voices, available in multiple languages and variants.
- **Azure's Text to Speech API**
  Azure's Text To Speech is a Speech service feature that converts text to lifelike speech with enable fluid, natural-sounding text to speech that matches the intonation and emotion of human voices. Available for more than 270 neural voices across 119 languages and variants.

When it comes to the other variant of the Speech modality, **Speech to Text** or transcription, the explored available options were:

- **Android's SpeechRecognizer API**
  The implementation of this API is likely to stream audio to remote servers to perform speech recognition. As such this API is not intended to be used for continuous recognition, which would consume a significant amount of battery and bandwidth.
- **Google Cloud's Speech-to-Text API**
  Google Cloud's Speech-to-Text API enables developers to convert audio to text in over 125 languages and variants, by applying powerful neural network models in an easy to

use API, powered by Google's AI technologies. Real-time speech recognition results are powered by Google's machine learning.

- **Azure's Speech to Text API**
  The Azure speech-to-text service analyzes audio in real-time or batch to transcribe the spoken word into text. Out of the box, speech to text utilizes a Universal Language Model as a base model that is trained with Microsoft-owned data and reflects commonly used spoken language. Customization options vary by language or locale.

B.2 Voice

Regarding **Voice Recording**, the explored technologies were:

- **Android's MediaRecorder API**
  MediaRecorder is a built-in class that helps to easily record video and audio files. It is Android's high-level framework for capturing audio and/or video. It records to a file directly, which can then be played back using MediaPlayer (covered later).
- **Android's AudioRecord API**
  The AudioRecord API manages the audio resources for Java applications to record audio from the audio input hardware of the platform. AudioRecord removes a layer of abstraction between the application and a device's audio hardware, recording uncompressed audio with no way to write directly to a file.

To perform **Voice Reproduction**, the following options were considered:

- **Android's MediaPlayer API**
  With MediaPlayer APIs it is possible to play audio or video from media files stored in application resources (raw resources), from independent files in the file system, or from a stream of data arriving over a network connection. Once initialized, MediaPlayer can be started, paused, and stopped, providing straightforward playback.
- **Android's AudioTrack API**
  The AudioTrack API manages and plays a single audio resource for Java applications. It allows streaming of PCM audio buffers to the audio sink for playback. It can be used to either stream audio continuously or play short sounds that fit in memory.

B.3 Images

When it came to **Sketch Recognition**, **ML Kit's Digital Ink Recognition API** was considered. Machine Learning Kit brings Google's machine learning expertise to mobile developers in a powerful and easy-to-use package. With ML Kit's Digital Ink Recognition API, it is possible to recognize handwritten text on a digital surface in hundreds of languages, as well as classify sketches.

To achieve a **Word to Image** implementation, the **ARASAAC REST API for pictograms** was considered. ARASAAC offers graphic resources and materials to facilitate communication and cognitive accessibility to all people who, due to different factors (autism,

intellectual disability, lack of knowledge of the language, elderly , aphasia, among others), present serious difficulties in these areas, which makes it difficult to include them in any area of daily life.

B.4 Camera

The following options were considered and explored regarding the **Object Detection** functionality:

- **TensorFlow Object Detection API**
  The TensorFlow Object Detection API is an open source framework built on top of TensorFlow that makes it easy to construct, train and deploy object detection models, detecting, locating and tracing an object from a still image or video. TensorFlow is a free and open-source software library for machine learning and artificial intelligence developed by Google.
- **ML Kit's Object Detection and Tracking API**
  ML Kit's Object Detection and Tracking API can detect and track objects in an image or live camera feed. When an image is passed to ML Kit, it detects up to five objects in the image along with the position of each object in the image. When detecting objects in video streams, each object has a unique ID that can be used to track the object from frame to frame. Also developed by Google.

B.5 Text

To make use of a **Translation** functionality, the following options were explored:

- **Google Cloud's Translation API**
  Translation API Basic uses Google's neural machine translation technology to instantly translate texts into more than one hundred languages. Translation API Advanced offers the same fast, dynamic results you get with Basic and additional customization features.
- **Azure's Text Translation REST API**
  Text Translation is a cloud-based feature of the Azure Translator service and is part of the Azure Cognitive Service family of REST APIs. The Text Translation API can translate text with a simple REST API call between language pairs across more than 100 languages and supported dialects.

B.6 Storage

Regarding the **Storage** services, the following options were explored:

- **Firebase Storage Services**
  Firebase works with two different databases, the Real-Time Database that is the original Firebase database product, and Cloud Firestore that is a new and improved version of the Real-Time Database. Cloud Storage for Firebase is a powerful, simple, and cost-effective object storage service built for Google scale.

- **Google Cloud Storage**
  Google Cloud Storage is a RESTful online file storage web service for storing and accessing data on Google Cloud Platform infrastructure. The service combines the performance and scalability of Google's cloud with advanced security and sharing capabilities.
- **Azure Storage**
  The Azure Storage platform is Microsoft's cloud storage solution for modern data storage scenarios. Azure Storage offers highly available, massively scalable, durable, and secure storage for a variety of data objects in the cloud.

APPENDIX C

First Developed version of the User Interfaces of Afluentia

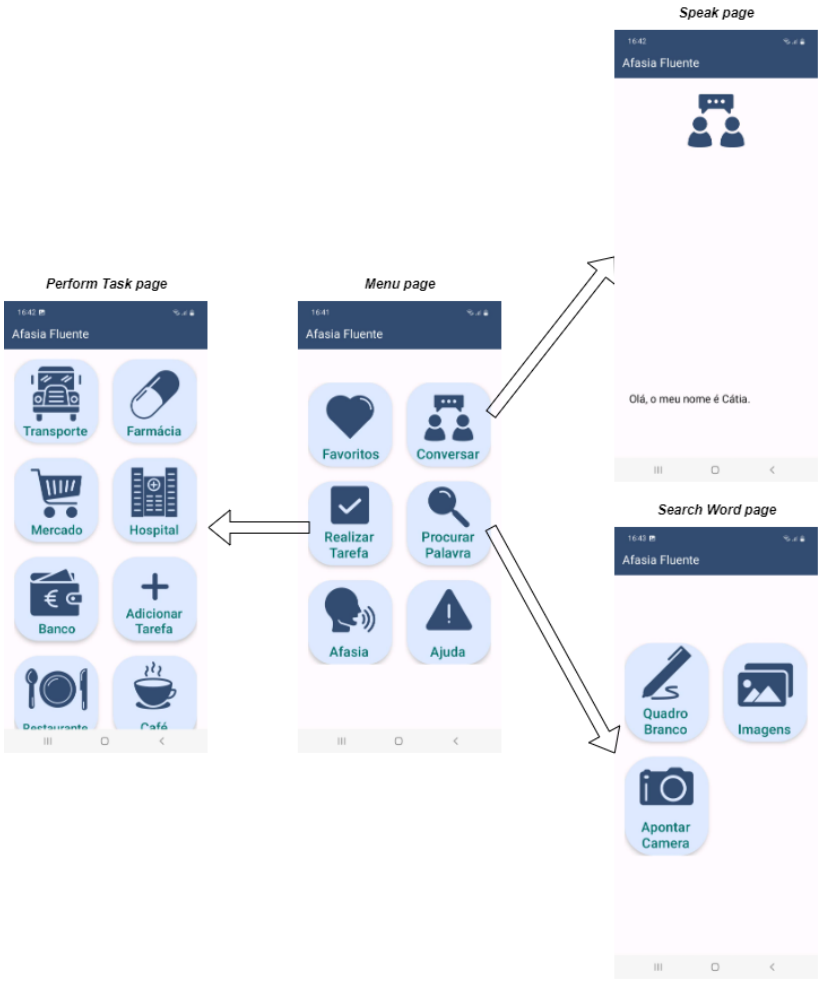


Figure C.1: First version of the main developed interfaces, that derive from the Menu page, being them the Speak page, Perform Task page and Search Word page.

The user interfaces corresponding to the first version of the main developed interfaces that derive from the Menu page, being them the Speak page, Perform Task page and Search Word page, are depicted in Figure C.1.

Figure C.2: First version of the pages deriving from the Search Word page, being them the White Board page, Camera page, Context for Words page, Fitzgerald Key page, Words page and Image for Word page.

The user interfaces corresponding to the first version of the pages deriving from the Search Word page, being them the White Board page, Camera page, Context for Words page, Fitzgerald Key page, Words page and Image for Word page, are depicted in Figure C.2.

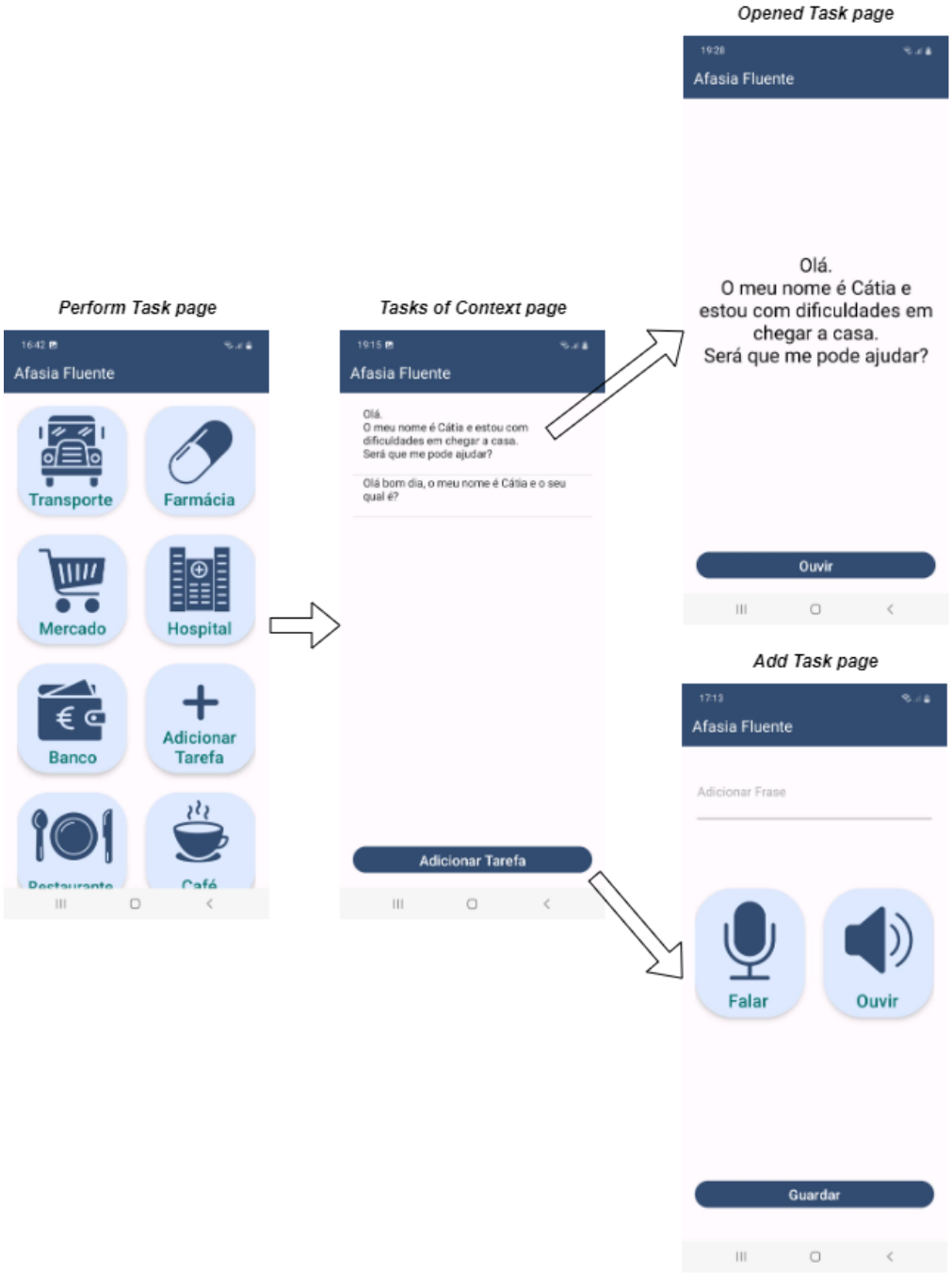


Figure C.3: First version of the pages deriving from the Perform Task page, being them the Tasks of Context page, Opened Task page and Add Task page.

The user interfaces corresponding to the first version of the pages deriving from the Perform Task page, being them the Tasks of Context page, Opened Task page and Add Task page, are depicted in Figure C.3.

APPENDIX D

Guide with Question and Tasks for Focus Groups

A aplicação Afluentia tem como principal objetivo servir como um assistente de comunicação multimodal a utilizadores diagnosticados com afasias do tipo fluente. Durante o processo de desenvolvimento da mesma, foram identificados três domínios principais, sendo eles Conversar, Realizar Tarefa e Procurar Palavra. No último domínio foram identificados três subdomínios, sendo eles Quadro Branco, Imagens e Apontar Câmera. Todos eles estão apresentados de seguida.

Para cada um dos domínios e/ou subdomínios foram identificadas tarefas a realizar. De modo a realizar as mesmas, e tendo em conta o conhecimento que se possa ter em relação a afasias do tipo fluente, o utilizador deve pôr-se no lugar de um utilizador diagnosticado com qualquer uma dessas condições.

Conversar: O principal objetivo do domínio Conversar é o utilizador conseguir acompanhar uma conversa em que esteja inserido, de modo a poder colmatar dificuldades de compreensão. A conversa é transcrita em tempo real e o sistema poderá dar indicações para abrandar o ritmo da conversa e/ou utilizar frases mais simples e curtas.

Tarefas:

1. *Estou a conversar com um amigo, mas estou a ter algumas dificuldades em acompanhar o que está a ser dito. Acompanhar a conversa através da aplicação vendo a sua transcrição.*

Realizar Tarefa: O principal objetivo do domínio Realizar Tarefa é o utilizador poder realizar tarefas no seu dia-a-dia tendo em conta diferentes contextos. O utilizador tem ao seu dispor alguns ficheiros escritos/aúdio e pode também adicionar os seus, tendo em conta as especificidades das suas necessidades.

Tarefas:

1. *Amanhã tenho de ir ao supermercado comprar ração para o meu cão. Como tenho receio de chegar lá e não conseguir encontrar a ração e pedir ajuda para esse fim, verificar se a aplicação já contém alguma função pré definida para este efeito.*
2. *Caso a aplicação não contenha o pretendido, adicionar conteúdo para esse fim de forma a conseguir perguntar "Onde está a comida para cão?".*
3. *Caso tenha adicionado um conteúdo por escrito, adicionar agora um conteúdo por voz. Ou vice-versa.*
4. *Sabendo da existência destas duas formas de adicionar conteúdo, considera alguma como mais provavelmente usada?*
5. *Aceder à tarefa que guardou.*

Procurar Palavra: O principal objetivo do domínio Procurar Palavra é o utilizador, através dos três subdomínios seguintes que tem ao seu dispor, conseguir chegar a uma ou mais palavras específicas.

Quadro Branco:

Através do subdomínio Quadro Branco o utilizador poderá desenhar e o sistema reconhecerá o que o utilizador desenhou, ou mesmo escrever e o sistema devolverá o que o utilizador possa ter escrito. Qualquer palavra que o sistema devolve, o utilizador tem a possibilidade de ouvir a mesma.

Tarefas:

1. *Quero dizer uma palavra específica (coração), mas não consigo verbalizar a mesma, apenas consigo lembrar-me da sua forma. Utilizar o sistema para desenhar o pretendido e obter assim a palavra reconhecida.*

Imagens:

Através do subdomínio Imagens, o utilizador tem ao seu dispor os mesmos contextos presentes em Realizar Tarefa. Abrindo um desses contextos, encontra palavras (Descrições (Adjetivos e Advérbios), Nomes, Verbos, Pessoas & Pronomes e Preposições & Social) tendo em conta o contexto inicial. O utilizador tem também a possibilidade de ouvir essa mesma palavra e de ver uma imagem associada à mesma.

Tarefas:

1. *Quero dizer uma palavra específica (conduzir), mas não consigo verbalizar a mesma, apenas consigo lembrar-me da sua primeira letra. Utilizar o sistema para introduzir a primeira letra, tendo em conta o seu contexto, e obter assim a palavra pretendida, tal como uma imagem associada à mesma.*

Apontar Câmera:

Através do subdomínio Apontar Câmera o utilizador pode apontar a câmera do seu telemóvel a qualquer objeto e o sistema irá identificar esse objeto em tempo real.

Tarefas:

1. *Tenho um objeto à minha frente (rato do computador), e não consigo verbalizar do que se trata. Utilizar o sistema para identificar esse mesmo objeto.*

APPENDIX E

Final Developed version of the User Interfaces of Afluentia

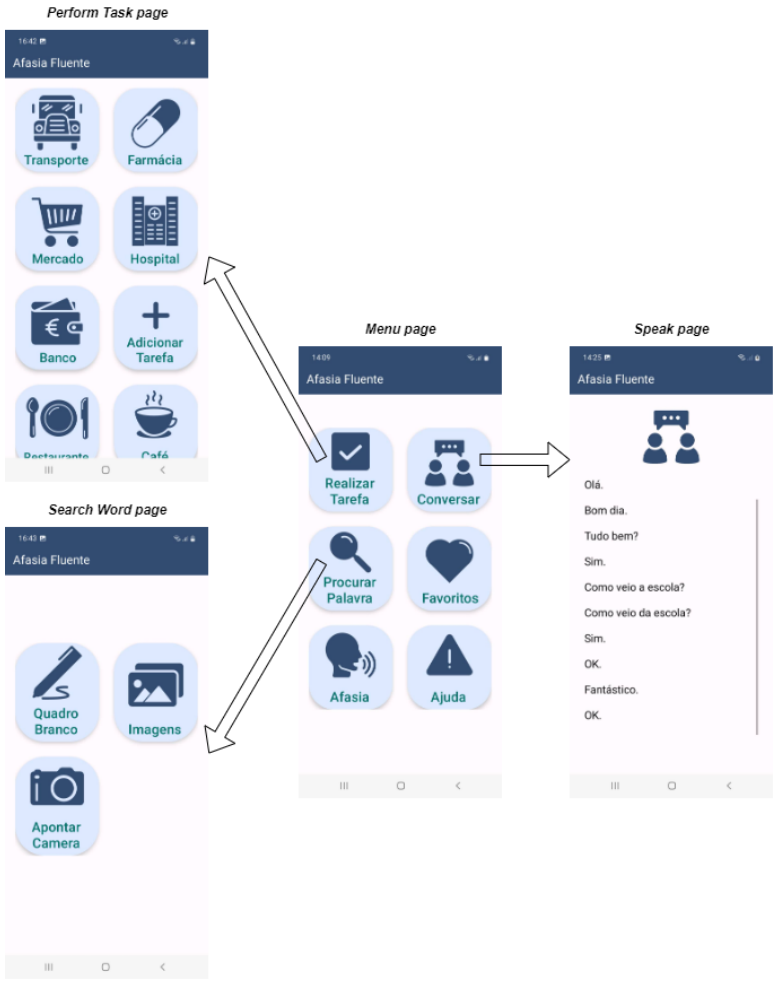


Figure E.1: Final version of the main developed interfaces, that derive from the Menu page, being them the Speak page, Perform Task page and Search Word page.

The user interfaces corresponding to the first version of the main developed interfaces that derive from the Menu page, being them the Speak page, Perform Task page and Search Word page, are depicted in Figure E.1.

Figure E.2: Final version of the pages deriving from the Search Word page, being them the White Board page, Camera page, Context for Words page, Words page and Image for Word page.

The user interfaces corresponding to the first version of the pages deriving from the Search Word page, being them the White Board page, Camera page, Context for Words page, Words page and Image for Word page, are depicted in Figure E.2.

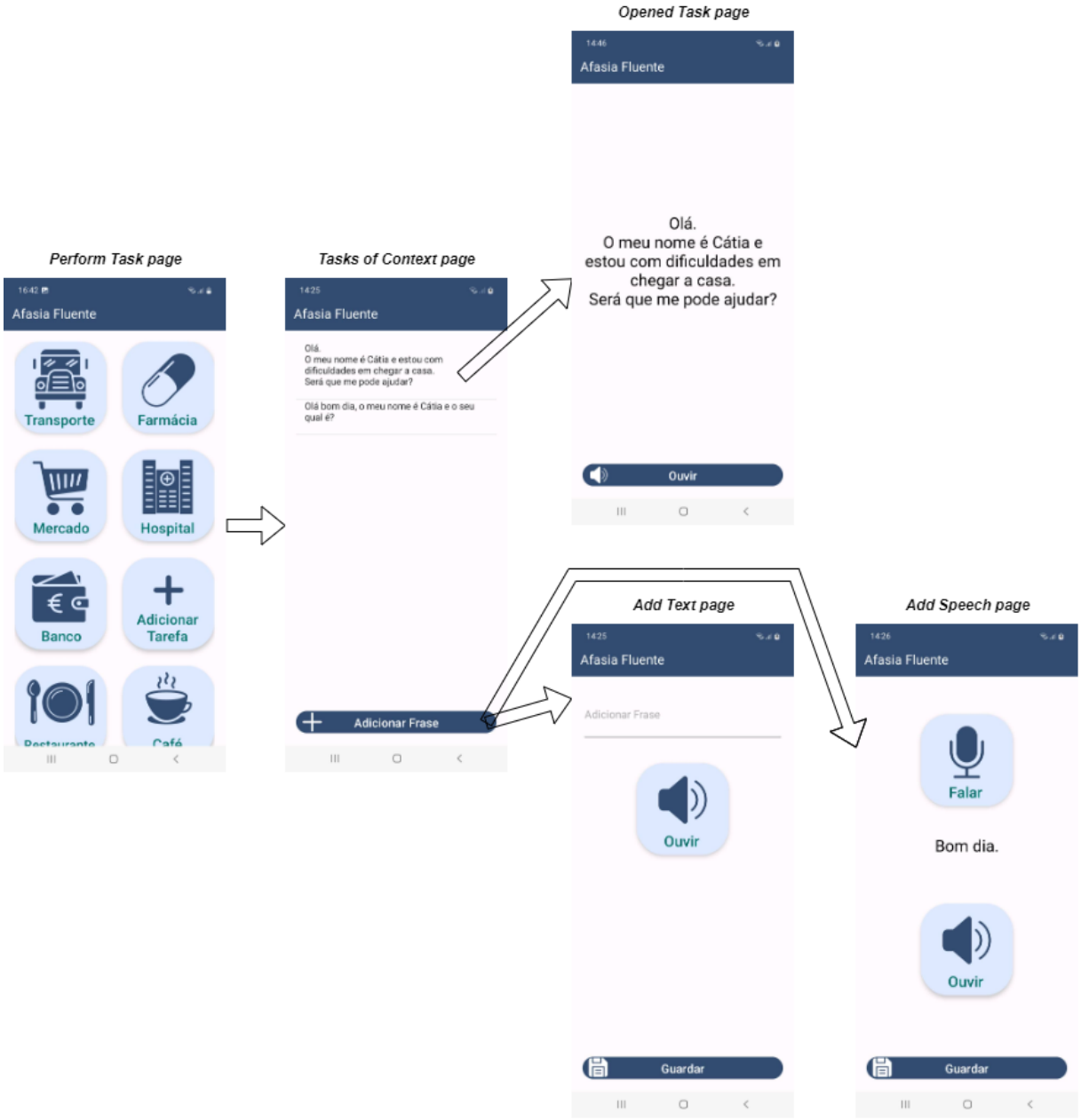


Figure E.3: Final version of the pages deriving from the Perform Task page, being them the Tasks of Context page, Opened Task, Add Text page and Add Speech page.

The user interfaces corresponding to the first version of the pages deriving from the Perform Task page, being them the Tasks of Context page, Opened Task, Add Text page and Add Speech page, are depicted in Figure E.3.

References

[1] M. Cruice, L. Worrall, L. Hickson, and R. Murison, "Finding a focus for quality of life with aphasia: Social and emotional health, and psychological well-being," *Aphasiology*, vol. 17, no. 4, pp. 333–353, 2003. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/02687030244000707.

[2] J. Fonseca, A. Dias, A. Rofes, A. C. Ferreira, A. Carvalhal, A. Matos, C. Maruta, C. Magalhães, D. Nogueira, D. Parente, F. Garcia, F. Miranda, G. Leal, G. Valido, I. T. Rodrigues, I. Fonseca, I. P. Martins, L. Taveira, M. Moreira, P. Vital, P. Pestana, R. Branco, R. Paixão, and V. de Aguiar, *Afasia e Comunicação após Lesão Cerebral*, 1. Papa-Letras, 2018, ISBN: 9789898214652.

[3] S. K. Kane, B. Linam-Church, K. Althoff, and D. McCall, "What we talk about: Designing a context-aware communication tool for people with aphasia," in *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, 2012, pp. 49–56. [Online]. Available: https://dl.acm.org/doi/10.1145/2384916.2384926.

[4] N. Simmons-Mackie, L. Worrall, C. Shiggins, J. Isaksen, R. McMenamin, T. Rose, Y. E. Guo, and S. J. Wallace, "Beyond the statistics: A research agenda in aphasia awareness," *Aphasiology*, vol. 34, no. 4, pp. 458–471, 2020. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/02687038.2019.1702847.

[5] C. Brandenburg, L. Worrall, A. D. Rodriguez, and D. Copland, "Mobile computing technology and aphasia: An integrated review of accessibility and potential uses," *Aphasiology*, vol. 27, no. 4, pp. 444–461, 2013. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/02687038.2013.772293.

[6] K. Moffatt, G. Pourshahid, and R. M. Baecker, "Augmentative and alternative communication devices for aphasia: The emerging role of "smart" mobile devices," *Universal Access in the Information Society*, vol. 16, no. 1, pp. 115–128, 2017. [Online]. Available: https://link.springer.com/article/10.1007/s10209-015-0428-x.

[7] J. McGrenere, R. Davies, L. Findlater, P. Graf, M. Klawe, K. Moffatt, B. Purves, and S. Yang, "Insights from the aphasia project: Designing technology for and with people who have aphasia," in *Proceedings of the 2003 conference on Universal usability*, 2002, pp. 112–118. [Online]. Available: https://dl.acm.org/doi/10.1145/957205.957225.

[8] A. C. Caldas, *A Herança de Franz Joseph Gall*, 1. McGrawHill, 1999, ISBN: 9789727730414.

[9] D. Phutela, "The importance of non-verbal communication," *IUP Journal of Soft Skills*, vol. 9, no. 4, p. 43, 2015. [Online]. Available: https://www.proquest.com/openview/52442af596bbd7cc0220950cc1a9a3f2/1?pq-origsite=gscholar&cbl=2029989.

[10] S. E. Jones and C. D. LeBaron, "Research on the relationship between verbal and nonverbal communication: Emerging integrations," *Journal of communication*, vol. 52, no. 3, p. 499, 2002. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1460-2466.2002.tb02559.x.

[11] A. R. Damasio, "Aphasia," *New England Journal of Medicine*, vol. 326, no. 8, p. 531, 1992. [Online]. Available: https://www.nejm.org/doi/full/10.1056/NEJM199202203260806.

[12] ——, *Acquired aphasia*. Academic Press San Diego, CA, 1998, vol. 3, pp. 25–41. DOI: https://doi.org/10.1016/B978-0-12-619322-0.X5000-3.

[13] V. Suri, *Clinical Neurological Examination and Localization*. Springer, 2021. DOI: https://doi.org/10.1007/978-981-16-1228-2.

[14] PsychDB. (2021). "Approach to aphasia," [Online]. Available: https://www.psychdb.com/neurology/approach-aphasia#approach-to-aphasia (visited on 08/04/2022).

[15] Y. Elsahar, S. Hu, K. Bouazza-Marouf, D. Kerr, and A. Mansor, "Augmentative and alternative communication (aac) advances: A review of configurations for individuals with a speech disability," *Sensors*, vol. 19, no. 8, 2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6515262/.

[16] B. T. Ogletree, *Augmentative and Alternative Communication*, 1. Plural Publishing Inc, 2021, ISBN: 9781635502862.

[17] J.-Y. Mao, K. Vredenburg, P. W. Smith, and T. Carey, "The state of user-centered design practice," *Communications of the ACM*, vol. 48, no. 3, p. 105, 2005. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/1047671.1047677.

[18] A. Cooper, R. Reimann, and D. Cronin, *About face 3: the essentials of interaction design*. John Wiley & Sons, 2007, pp. 75–123, ISBN: 9780470084113.

[19] M. Maing, "Physical or virtual?: Effectiveness of virtual mockups of building envelope systems," in *BEST3 Conference*, 2012. [Online]. Available: https://www.brikbase.org/content/physical-or-virtual-effectiveness-virtual-mockups-building-envelope-systems.

[20] M. Walker, L. Takayama, and J. A. Landay, "High-fidelity or low-fidelity, paper or computer? choosing attributes when testing web prototypes," in *Proceedings of the human factors and ergonomics society annual meeting*, Sage Publications Sage CA: Los Angeles, CA, vol. 46, 2002, pp. 661–665. [Online]. Available: https://journals.sagepub.com/doi/abs/10.1177/154193120204600513.

[21] J. Rudd, K. Stern, and S. Isensee, "Low vs. high-fidelity prototyping debate," *interactions*, vol. 3, no. 1, pp. 76–85, 1996. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/223500.223514.

[22] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 1990, pp. 249–256. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/97243.97281.

[23] J. Nielsen, "Enhancing the explanatory power of usability heuristics," in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 1994, pp. 152–158. [Online]. Available: https://dl.acm.org/doi/10.1145/191666.191729.

[24] T. Zuk, L. Schlesier, P. Neumann, M. S. Hancock, and S. Carpendale, "Heuristics for information visualization evaluation," in *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, 2006, pp. 1–6. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/1168149.1168162.

[25] J. Lumsden, *Handbook of research on user interface design and evaluation for mobile technology*. IGI global, 2008, vol. 1, ISBN: 9781599048727.

[26] C. Pope and D. Allen, "Observational methods," *Qualitative research in health care*, pp. 67–81, 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119410867.ch6.

[27] J. Kitzinger, "Qualitative research. introducing focus groups.," *BMJ (Clinical Research ed.)*, vol. 311, no. 7000, pp. 299–302, 1995. [Online]. Available: https://europepmc.org/article/med/7633241.

[28] J. R. Lewis, "The system usability scale: Past, present, and future," *International Journal of Human–Computer Interaction*, vol. 34, no. 7, pp. 577–590, 2018. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/10447318.2018.1455307.

[29] A. A. Al-Hassan, B. AlGhannam, M. B. Naser, and H. Alabdulrazzaq, "An arabic translation of the computer system usability questionnaire (csuq) with psychometric evaluation using kuwait university portal," *International Journal of Human–Computer Interaction*, vol. 37, no. 20, pp. 1981–1988, 2021. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/10447318.2021.1926117.

[30] X. Du, M. Dai, H. Tang, J.-L. Hung, H. Li, and J. Zheng, "A multimodal analysis of college students' collaborative problem solving in virtual experimentation activities: A perspective of cognitive load," *Journal of Computing in Higher Education*, pp. 1–24, 2022. [Online]. Available: https://link.springer.com/article/10.1007/s12528-022-09311-8.

[31] StatCounter. (2022). "Mobile operating system market share in portugal - june 2022," [Online]. Available: https://gs.statcounter.com/os-market-share/mobile/portugal.

[32] S. Bose *et al.*, "A comparative study: Java vs kotlin programming in android application development," *International Journal of Advanced Research in Computer Science*, vol. 9, no. 3, pp. 41–45, 2018.

[33] D. Paolieri and A. Marful, "Norms for a pictographic system: The aragonese portal of augmentative/alternative communication (arasaac) system," *Frontiers in psychology*, vol. 9, 2018. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02538/full.