




Making social media applications inclusive for deaf end-users with access to sign language

Alexandros Yeratziotis¹  · Achilleas Achilleos² · Stavroulla Koumou³ · George Zampas³ · Regan A. Thibodeau⁴ · George Geratziotis¹ · George A. Papadopoulos³ · Iasonos Iasonas² · Christophoros Kronis²

Received: 21 February 2023 / Revised: 2 August 2023 / Accepted: 19 September 2023 /
Published online: 10 November 2023
© The Author(s) 2023

Abstract

Social media apps such as WhatsApp, Messenger, Telegram and Viber are primary communication channels for most people today. Even within these Mainstream Social Media Applications, persons who are deaf are not provided with the rights and means to interact using sign language. This work, an outcome of the research project ‘‘Accessible System and Social Media Mobile Application for Deaf Users (ASM4Deaf)’’, supports the inclusion of deaf end-users in social media applications. This was achieved via: 1) design and development of the ASM4Deaf system, which includes the Connect Deaf mobile application, enabling use of Sign Language Alphabet keyboards in social media apps in 17 different sign languages and 2) the evaluations of Low- and High-Fidelity prototypes aimed to enhance the app’s design and functionality, i.e. the ability to browse, search and edit animated videos/GIFs in the American Sign Language (ASL), using the face swapping feature. This makes the full set of features offered by social media applications accessible to deaf end-users and their personal network, which is the main contribution of this work. A Human-Centred Design methodology was employed, with the end-users at the heart of the process, to design and evaluate the Lo-Fi and Hi-Fi prototypes, and based on the evaluation results to develop the ASM4Deaf system and the Connect Deaf mobile application to fully meet the requirements of the deaf end-users.

Keywords Inclusive design · Accessibility · Deaf end-user · Face swapping · Mobile application · Sign Language Alphabet · Human-Centred Design

1 Introduction

Population statistics for individuals who are deaf have not been accurately recorded to date due to deficiencies in demographic sampling surveys. In spite of that, there are estimates for the size of deaf and hard-of-hearing populations, mainly provided through the Bureau of Census, deaf organizations and associations, and general deaf statistics. The World Health Organization (WHO) estimates that there are around 430 million people worldwide

at present experiencing hearing loss [1]. This refers to hearing loss that is greater than 40 dB in the better hearing ear of adults and greater than 30 dB in the better hearing ear of children. From this overall total, the World Federation of the Deaf (WFD) estimates that there are approximately 70 million people worldwide who are deaf and use sign language as a first language, and there are more than 300 sign languages in use worldwide [2]. The European Union of the Deaf likewise provides information on its member countries and affiliated members, including estimates of sign language users and interpreters. Its members account for over 1 million sign language users [3]. In the United States, American Sign Language (ASL) is the primary language of communication for over 500,000 people [4].

Most individuals who are deaf use a national sign language as a first language, leading to up to 80% of them not being able to successfully understand written content [5]. This literacy barrier effect on their lives is multifaceted and consequently negative. In the context of this research, we explored the effect from the facet of Information Communication Technologies (ICTs) experiences, especially in Mainstream Social Media Applications (MSMAs).

With the aforementioned in mind, we envisioned realising a solution that could contribute and impact on multiple levels. Firstly, technology-based innovation should be utilised to increase access, awareness and support for effective, yet entertaining communication in local sign language. It should exploit a broad range of ICTs and video media in achieving this impact. Secondly, it should impact on learning and improving literacy skills of sign language users, through the provision of inclusive and equal access in MSMAs. This includes gender and diversity inclusive innovation. Thirdly, it should be replicable and adaptable to country contexts, meaning that it can be reproduced in different sign languages. Fourthly, it should be an affordable innovation, making it scalable in the community. Fifth, it must be a usable and useful innovation for persons who are deaf and their personal network (including hearing), i.e. family members, colleagues, friends, educators and/or other community members, who engage in learning and communicating in sign language. Sixth, its design must be founded on the principles of Human-Centred Design (HCD). In conclusion, the solution needed to be innovative, to enable the creation of content and improve access to communication and education in local sign language, act as a literacy intervention for persons who are deaf and contribute to the research area of Accessible Computing.

The ASM4Deaf funded project aimed to develop a cloud-based system and extend the Connect Deaf innovative mobile application to offer the full set of features, i.e. to support the use of multiple sign languages within MSMAs, including WhatsApp, FB Messenger, Google Hangouts, Viber, Telegram, etc. Moreover, it should allow using ASL interpreter videos/GIFs, based on a repository of phrases/words created in the project, utilising the face-swapping feature and posting the customised video to social media. Being typically designed with those that hear in mind, people who are deaf are not provided with the rights and means to interact and communicate with contacts in national sign language within these popular social media apps. This work supports the inclusion of the deaf community in MSMAs.

In this paper, we present the Hi-Fi and functional prototypes of the Connect Deaf mobile application and report on findings from focus groups conducted with end-users in Cyprus and the US, who had evaluated the Low-Fidelity (Lo-Fi) and High-Fidelity (Hi-Fi) prototypes respectively. Findings from the focus groups conducted with end-users in Cyprus to evaluate the Lo-Fi prototype and the prototype itself were presented in [6] and are therefore only summarised in this paper. The front-end, i.e. the content management system, and

the face swapping approach implemented within the mobile application are discussed. The ASM4Deaf system's architecture and back-end were likewise presented in [6]. The paper consists of six sections; Sect. 2 discusses related work, Sect. 3 describes the methodology, Sect. 4 presents the ASM4Deaf system from a technical perspective. The Hi-Fi prototype follows in Sect. 5 and the results from the focus groups are discussed in Sect. 6. The paper closes with conclusions and future research perspectives in Sect. 7.

2 Related work

Yeratziotis and Van Greunen [7] stressed that access to information and the ability to share information online, must not be a privilege for abled end-users only. Making ICTs accessible for persons with disabilities must be a priority, as it can contribute toward reducing the digital divide and supporting inclusion. To accomplish this, one must keep in mind not only end-users' disabilities, but also their abilities. Topics such as User Interface (UI) design for deaf end-users, e-learning multimedia, sign language and literacy barriers, deaf end-users experience, were explored. The primary outcome of the work was the identification of User Experience (UX) and UI design guidelines for the design of mobile applications for end-users who are deaf. The resulting guidelines were then applied to design the mobile application prototype "SignChat", which was evaluated with deaf end-users. The application's goal would be to provide a more cost effective communication method by using Sign Language Alphabet (SLA) keyboards and to address the existing communication barrier between spoken language and sign language end-users through message translations.

A Heuristic Evaluation for the Deaf Web User Experience (HE4DWUX) was proposed in [8]. This was a novel usability inspection method to assist experts in the field of Human-Computer Interaction (HCI) and Web developers alike, when evaluating and designing websites for end-users who are deaf. This was performed by measuring accessibility and usability problems that can influence their UX on the Web. In terms of solutions, recent noteworthy ones, which are focused on increasing access to sign language include Studio KSL, World Around You and Señas y Sonrisas (Signs and Smiles). Studio KSL is a platform that documents Kenyan Sign Language (KSL) in a visual glossary and produces KSL videos for integration into accessible books in Kenya. World Around You is a platform that documents, collects and shares local sign and written languages in the Philippines. Señas y Sonrisas (Signs and Smiles) documents a corpus of Nicaraguan Sign Language (NSL) and has an associated language learning mobile app with downloadable lessons, as well as a literacy outreach program to train parents of deaf children in Nicaragua [9].

The importance of a definitive description of published manuscripts in the field of Interactive Software Technology (IST) for users who are Deaf in the discipline of HCI focusing on accessibility is presented in [10]. It aimed to contribute in classifying topics for First Stage Researchers (R1) and Recognised Researchers (R2) entering the field. The authors: i) constructed a map of existing research topics in the field of IST, for deaf end-users, in the discipline of HCI focusing on accessibility, ii) summarized the purpose of each code category of the map, and iii) identified the least and most researched topics of the map thereof.

The lack of support and marginalisation of Cypriot deaf people, due to lack of technological and financial support, are elaborated in [11]. Subsequently, the communication needs and requirements of Cypriot deaf people were further explored implementing

user-centred design methods. This led to the design of a functional prototype of a mobile app to help them communicate more effectively with hearing people. Although it is mentioned that more experiments are needed in order to test the performance of the mobile app under specific conditions, such as in noisy environments, and other functional issues impacting on its usefulness, the app showed real promise. Its potential, in reducing the need for deaf persons to rely on the presence of someone to act as an interpreter to support their completion of everyday activities was especially compelling. In terms of its potential, evaluation results also emphasise that it can support deaf people in Cyprus to engage more effectively in communication within society and achieve social acceptance.

Communication and interaction between people are fundamentally accomplished with the use of languages, which are a necessary tool for the society to socially interact and make significant progress. Hence, it is vital to promote social awareness towards the understanding of people with hearing disabilities, favoring communication between hearing and deaf people. Accessibility and social inclusion are important issues when it comes to language for deaf end-users and the ability to communicate with other people, since learning sign language is not an easy task. This is due to different reasons, such as access to sign languages information being scarce, not having interpreters of this language or because learning of this language is not taught from childhood. The work performed in [12] designed and developed a mobile application with augmented reality. The objective and contribution of this work was to determine the influence it has on learning the Peruvian sign language (PSL). The mobile application was evaluated with a sample of 30 students as a control group and an experimental group. In accordance to the results obtained, the students started with a performance percentage of 23%, demonstrating an important change when using the mobile application, significantly improving user performance to 65% with respect to learning the PSL.

Deaf end-users prefer to utilize their first language, i.e. sign language, to perform daily activities. This is specifically, due to the fact that communication is greatly affected by disorders such as hearing and speech impairment. Several assistive technologies and tools have been developed that can help deaf and mute end-users such as speech-to-text, speech-to visual and sign language. The focus of this research performed in [13] is to develop a mobile application to break the barrier between hearing and non-hearing or mute in terms of communication. This is done by assisting the mute or deaf people in the daily activities of whether it be used inside of a household or outside, as well as the communication with other people without disabilities. The mobile application offers four modules. The Communication module has four (4) sub-modules: 1) The Text-to-Speech sub-module lets the user input word/phrases and reads it out loud, 2) The Speech-to-Text sub-module allows voice input and the user can convert it to fingerspelling sign language, 3) The FingerSpelling sub-module accepts text input and can convert it to fingerspelling sign language; i.e. both in American Sign Language (ASL) and Filipino Sign Language (FSL) and 4) The Scenarios sub-module has four (4) categories and contains a dictionary of words that is usually used in each place, which when clicked shows an image of the word in the FSL. The Alert module can be used by the deaf end-user to draw the attention of people around them, e.g. in case they are in danger. The Image Dictionary module contains the fingerspelling arranged alphabetically. Finally, the Help module provides instructions to the end-user on how to use the application. The mobile application was tested, analysed and improved as the researchers gathered more data, conducted interviews in communities and with pathologists, and analyzed that data to further enhance the application for a better output.

Worthy initiatives over the past twenty years for improving accessibility and inclusion on the Web and ICTs for deaf end-users are the Web Content Accessibility Guidelines

(WCAG) of the W3C [14], DictaSign [15, 16], ViSiCAST [17], eSign [18], SignLinkStudio [19], the European Echo Project [20], SignStream [21, 22], the University of Bristol's British Sign Language (BSL) Moodle system [23], Ohio State University's digital storytelling system [23], the South African Sign Language (SASL) machine translation project of Stellenbosch University [24] and Paula of DePaul University [25].

Efforts to make ICTs more inclusive to deaf end-users are imperative, while the need for more research contributions in this area is likewise stressed [5]. More inclusive and accessible Web and ICTs have a significant influence on the lifestyle of deaf persons in many contexts, among other; education, employment, access to information and services, entertainment and social [5].

Next, face swapping is discussed, an important technical approach implemented in the development of the updated version of the Connect Deaf app for the ASM4Deaf system. This will then lead into a comparison of similar applications in Sect. 2.2, thus providing an overview of the current state of technology in this area.

2.1 Face swapping

Gaining its fame mainly from the face/head swap Snapchat filter, Face Swapping is nowadays used in a variety of applications, mainly in the domain of entertainment [9, 26]. Like in Snapchat, it is often offered as a one-of-the-many filter in a collection, but it can also be found in dedicated entertainment applications, like Reface [27]. Additional application prospects of the Face Swap technology are located in the domain of privacy protection [28] and the theatrical industry [29, 30].

Face Swapping is one of the Artificial Intelligence-synthesized content categories known as deep fakes (“deep learning” + “fake”), along with Lip-Syncing and Puppet-Mastering [31]. The term Face Swapping refers to the replacement of the face-area in a data source A with the face-area from a data source B, retaining the rest of information of data source A as is. A data source involved in such a process can be either an image or a video file of/with the face of an individual. This is so that, in essence, an identity replacement/transfer is possible to be carried out between two such sources [30]. Some well-known examples of identity replacement are the scene from Home Alone with Macaulay Culkin's face swapped with Sylvester Stallone's in 2019 and the DeepTomCruise TikTok account that attracted the attention of the audience with high-quality deep fake videos in 2021 [32].

For the detection and/or replacement of the face-area in the images/videos used in Face Swapping both face landmark detection algorithms from the Computer Vision domain and models from the Machine Learning and Artificial Intelligence domains are currently in use. Given the extensive literature and matching test code-repositories as of today available, the research communities involved do seem quite engaged. Proposing new algorithms [33, 34] and developing new models [31] for improving result quality, optimizing performance (execution time and/or memory) [30], exploring the application prospects of the Face Swapping technique [9, 26, 28, 29], as well as prevention of biometric-based security breaches it may be used for [35, 36], are only a few of the topics of interest.

2.2 Relevant apps

From a review of the literature and marketplace different mobile applications have been identified as listed in Table 1, which are separate applications that needed to be installed in Android and iOS. The main need that was identified was to develop a mobile application

Table 1 Comparison of similar apps

Work	Type	Description	AI	Other Features	Status	Cost
SLAIT ²	Mobile app	Real-time ASL to text and speech to text translation	Yes	Video communication	Beta version	Not in the market yet
Hand Talk Translator ³	Mobile app	Real-time text and audio to ASL and Brazilian Sign Language	Yes	Avatar as a signer	Available	Free to download with in-app purchases (from €2.49 for 10 coins)
Connect Deaf ⁴ (version 1)	Mobile app	SLA keyboards for 13 countries	No	Posting of SLA GIFs and images and Sign Language stickers and emoticons to MSMAs	Available	€1.99
Signily ⁵	Mobile app	ASL keyword	No	Posting of (color) customized SLA messages	Retired since 2020	€0.99
“ASL to English” developed by Priyanjali Gupta ⁶	GitHub project	Real-life ASL to text translation, only for six specific sign-language words and phrases	Yes	-	Available	Free

²SLAIT website: <https://slait.ai/>³Hand talk website: <https://www.handtalk.me/en/app/>⁴Connect Deaf (Android, version 1): <https://play.google.com/store/apps/details?id=com.mdl.signkeyboard>⁵ASLIZED! website: <http://aslized.org/innovations/>⁶“ASL to English” GitHub repository: https://github.com/priyaanjaali0611/ASL_to_English

that functions as a keyboard service within MSMA. This makes the full set of features offered by MSMA accessible to deaf end-users and their personal network, which is the main contribution of this work.

The Connect Deaf application (from hereafter Connect Deaf app version 1) was already in existence as an isolated application. It was designed and developed on the basis of the limited number of related applications, highlighting the need for more work in this area. It was a significant stride towards inclusiveness and accessibility for deaf end-users, enabling them to interact and communicate in their first language, making the experience more natural to them.

As part of the ASM4Deaf system, it was required to upgrade the Connect Deaf app version 1 from several perspectives: innovation, features and functionality. In its first release, the Connect Deaf app version 1 supported the use of SLA keyboards in MSMA for 13 national sign languages.¹ It thus provided, on Android and iOS, keyboard services for MSMA, benefitting deaf end-users and their personal networks.

In its evolution, as part of the ASM4Deaf system, the Connect Deaf application (hereafter Connect Deaf app version 2) has reached a level where it forms part of a cloud-based system, supporting 17 national sign language keyboards in 2D that can be used in MSMA, including Viber (not supported in the first release). In addition to posting SLA messages as GIFs and images and sharing sign language stickers and emoticons, designed and developed in the first release, it now also supports end-users in constructing their sign language videos/GIFs, at this stage in ASL, and implementing face swapping on the GIFs.

End-users have the ability to browse, search and edit animated videos/GIFs in ASL. Two keyboards are also available in both 2D and 3D, i.e. American and South African. All keyboards are offered in left-handed design too, a feature not available in the first release. Note, that this applies to one-handed SLAs. In its latest version, the Connect Deaf app version 2, aims to support sign language users in using their first language in their online chats in a more social, entertaining and educational manner. From a review of the literature and marketplace, similar apps with the Connect Deaf app version 1 were discovered, and are presented in Table 1.

Table 1 shows the comparison of the Connect Deaf app (version 1) with similar applications. The key difference of the Connect Deaf app is that it is not a separate mobile application, but it is actually a keyboard service application. This means that the user can install this application, enable it in the mobile phone and then be able to select it as the keyboard service to be used within the MSMA on Android and iOS. Moreover, version 2 of the mobile application, released as part of the ASM4Deaf system, offers the full set of features in order to facilitate inclusiveness and accessibility in MSMA for deaf end-users. This provides them with the capability to use and exploit all communication mechanisms in MSMA, as it is offered for hearing users.

3 Methodology

A HCD methodology was followed, based heavily on a co-creation approach. In HCD, a group of methods and principles are used that aim to support the design of useful, usable, pleasurable and meaningful products or services for people [37]. Four main activities compose the HCD methodology and to initiate the HCD activities, the need to design a new

¹ National sign languages in first release of Connect Deaf app: <https://connectdeaf.com/our-portfolio/connect-deaf/>

product or service must first be established. In the case of ASM4Deaf, the need was established, thus supporting the initiation of the HCD activities. There is a clear need to make MSMAAs more inclusive for deaf end-users, the key contribution of this work. It should be noted that a deaf end-user does have the option to record themselves signing a message with the camera of their smartphone and then posting it to their chat. However, this is based on the function of being able to post a video to the chat stream and was not designed with the deaf community in mind. It is a feature available to all end-users. Considering this, the option for sign language use in any form within the construction of messages is not offered by default in MSMAAs. Additionally, it was through discussions with members of the deaf community that led to shaping the specific idea in its current form, and determining its need thereof. Once the need for HCD is established, the following steps can commence [38]:

1. Understand and specify the context of use. Context of use encompasses the characteristics of the intended users, the tasks that the users will need to perform and the environment (organisational, technical and physical) in which the product or service is used. In ASM4Deaf the context of use is the mobile platform environment and the end-users utilise American Sign Language as a first language for communication purposes, as well as the majority have poor literacy skills in comparison to end-users who utilise a spoken language as a first language.
2. Specify the user and organisational requirements. This activity determines and specifies the major requirements for the new product or service. It needs to be extended by creating an explicit statement of user requirements. These requirements are dependent on the intended users, context of use and the organisational objectives. In ASM4Deaf, basic user requirements (see Sect. 4.1) were defined at the onset, and based on the feedback collected from end-users and experts during the co-creation and evaluation activities, these were finalised. Data collection methods included focus groups, prototyping and user surveys.
3. Produce design solutions. Potential design solutions are produced in this activity. Solutions are based on the description of context of use, results from any baseline evaluations and the established state of the art in the application domain. Design and usability standards and guidelines, as well as the experience and knowledge of the multidisciplinary design team are all crucial. Iteration is essential at this point and can result in additional user requirements. In ASM4Deaf, both Lo-Fi and Hi-Fi prototypes were designed and evaluated, resulting in several iterations in the design phase. In this activity, the co-design approach was particularly evident.
4. Evaluate designs against requirements. User-centred evaluation is necessary to determine whether the HCD methodology is successful. Moreover, new information on user requirements may surface and baselines can be established for comparing alternative designs. Feedback collected, quantitative and/or qualitative, can be used to progress a preferred design and assess its fulfillment of user and organisational objectives. Additionally, it helps monitor long-term use. In ASM4Deaf, the prototypes designed were useful tools in evaluations, guiding throughout whether the requirements of end-users have been satisfied.

Quantitative and qualitative data was collected from the end-users. Qualitative feedback is valuable, since the aim is centred on the inclusion of deaf end-users in the MSMAAs. As a group that is often overlooked in technology, and relatively closed towards the outside hearing community, trust needs to be gained. This was achieved with the assistance of a deaf consultant and sign language interpreters. Moreover, the end-users need to be actively involved in the process of designing and developing the product, if they are to use it. Next,

we will discuss the methods that were used for data collection and report on the inclusion and exclusion criteria in the recruitment of participants.

3.1 Prototyping

Lo- and Hi-Fi prototypes were designed and evaluated with the aim to collect user feedback and improve the design using a HCD methodology. A number of ideas were explored, some of which were more technically complicated to implement from the onset. It served as an opportunity, foremost to collect feedback on such ideas with potential to be implemented in the future, should the required desire be expressed by the end-users while addressing the current needs and requirements of the users via iterations. The Lo-Fi prototype was presented in [6]. The Hi-Fi prototype is presented in Sect. 5.

3.2 Focus group

Focus groups help acquire attitudes, reactions and opinions towards a product, as well as ideas, while at the same time aid in better understanding user requirements. Useful, in assessing end-user needs and feelings before and after implementation. Nielsen [39] suggests that it is more effective when the group consists of six to nine users and the session lasts for about two hours. The session will also need to be administered by a moderator, who maintains the focus of interests. However, there still needs to be a free-flowing and relatively unstructured style to the session. It is risky to use them as the only method when evaluating UIs; they can possibly produce inaccurate data, as users may think that they require one thing, when instead they require another. It is recommended to use them with other methods, like prototypes and scenarios. Focus groups are not only used to assess the usability of a design, but are also used to discover what it is that the end-users expect from the product. Four different focus groups were organised to evaluate the following prototypes:

1. Lo-Fi prototype with deaf end-users in Cyprus.
2. Lo-Fi prototype with deaf end-users in the US.
3. Hi-Fi prototype with deaf end-users in Cyprus.
4. Hi-Fi prototype with deaf end-users in the US.

In the context of this paper, we report on findings from focus groups 2, 3 and 4. Focus group 1 findings were reported in [6]. Sign language interpreters were present at all focus group sessions. For focus groups 1 and 3 a Cypriot sign language interpreter was present. For focus groups 2 and 4 an American sign language interpreter was present. Their role is invaluable, linking the research and technical team with the end-users of the respective deaf communities.

3.3 User survey

The System Usability Scale (SUS) was the evaluation instrument used. It provides a quick, easy to administer and at the same time reliable tool for measuring usability.² It consists of

² System Usability Scale (SUS): <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

a 10-item questionnaire with five response options for respondents; from Strongly agree to Strongly disagree. The SUS instrument was developed by John Brooke in 1986. It allows evaluating a wide variety of products and services, including hardware, software, mobile devices, websites and applications. In this work, the SUS wording was adapted in order to fit the evaluation of a mobile application. The items of the adapted SUS questionnaire are as follows:

- x1. I would like to use the app frequently
- x2. The app looks too complex
- x3. The app looks easy to u
- x4. I would need help to use the app
- x5. The different functions look well connected
- x6. The design of the screens looks similar
- x7. I would learn to use the app very quickly
- x8. The app looks difficult to use
- x9. I would be confident using the app
- x10. I would need to learn a lot of things before launching the app

The above items were video recorded in the American and Cypriot sign languages respectively. The videos were then integrated into the Google Form Questionnaire for each focus group. Items were presented in Greek text, supported with Cypriot sign language videos for deaf end-users in Cyprus³ and English text, supported with American sign language videos for deaf end-users in the US.⁴ The SUS instrument score is calculated using a formula as described here.⁵ Based on existing research performed on the definition and use of SUS, a score of above 68 would be considered above average and anything below 68 is below average. The general guideline on the interpretation of SUS score is also available in¹⁰ for interested readers.

3.4 Participant sampling

Criteria were determined in sampling end-users to participate in the study. The inclusion criteria were:

1. Teenage participant aged 15–18 or adult participant aged 18 and older.
2. American and/or Cypriot sign language user.
3. Owner of an iOS or Android smartphone.
4. User of a MSMA (e.g. Telegram, WhatsApp, FB messenger, Google Hangouts etc.).
5. Diversity in subjects in terms of ethnicity and gender.
6. From the hearing personal network of a deaf end-user (e.g. relative, friend, colleague) and not a user of ASL/CSL.

³ <https://forms.gle/Cx7UxTuiiUnSP2Pa8>

⁴ <https://forms.gle/MHAWsfxK2dYajGb8>

⁵ Interpreting System Usability Scale (SUS) Score:
<https://uiuxtrend.com/measuring-system-usability-scale-sus/>

In addition to the inclusion criteria, several exclusion criteria were also defined and included:

1. Deafblind persons.
2. No voluntariness.
3. No consent.

4 ASM4Deaf system

4.1 Baseline requirements

Table 2 describes the baseline requirements of the ASM4Deaf system that also defines the extensions to the Connect Deaf mobile app (version 2).

4.2 Architecture

The ASM4Deaf system architecture is composed by two main parts: 1) the frontend that will include the mobile application that the user will interact with and the platform that the administrator uses for uploading videos and images, and 2) the backend that will include the Web APIs for invoking from the mobile application the selected Face Swapping processing functionality and returning the final GIF back to the application. The backend also includes the web platform that enables the administrator to upload the original ASL GIFs and add relevant information (e.g. keywords), as well as manage these GIFs and their details. The GIFs and their information are stored respectively in the filesystem and a MySQL database. The content management system (CMS), implemented as one of the Python modules of the web platform, allows managing the GIFs, whereas the information of the GIFs will allow searching from the mobile application for specific GIFs, via the relevant Web APIs that will be implemented on the backend and invoked from the application.

Table 2 Baseline requirements defined at the onset

Sign language	American
Recorded GIF words	1000
Recorded GIF phrases	100
Face images	50
Interpreters in recordings	3 (minimum) interpreters to be used. Diversity in terms of ethnicity and gender are important requirements
2D SLA keyboards	4 new SLA keyboards to be designed and developed for the following countries: Russia, Canada, Mexico and Italy
Themes	5 new themes to be designed and developed
2D left-handed SLA keyboards	Connect Deaf app version 1 only offered right-handed keyboards for those based on using one hand when signing
	9 Left-handed SLA keyboards to be designed for the following countries: Cyprus, Greece, South Africa, America, Brazil, China, Germany, Ireland and France
3D SLA keyboards	2 3D keyboards to be designed and developed. For ASL and SASL

The aforementioned description represents the high-level architecture of the ASM4Deaf system.

The end-user through the smartphone is using the Connect Deaf mobile app version 2 (i.e. keyboard service) within the major social media applications (e.g. WhatsApp, Messenger) that offers the capability to type in the preferred SLA, while at the same time it can browse or search (e.g. by keyword—“Good morning”) and then select a video/GIF. In the case of the browse action an HTTP GET request is sent to invoke the appropriate Web API endpoint, to retrieve all videos/GIFs and present them to the end-user in a list. In the case of the search action an HTTP GET request is sent with the input data (i.e. keywords), so that the relevant videos/GIFs are returned by the invoked Web API endpoint. The end-user then selects the video/GIF and a face or emoticon from the available ones in the mobile application or takes a picture of his/her face and an HTTP POST request is sent to the relevant Web API endpoint including the face and video/GIF to be processed and transformed by the second Python module of the web platform, the GIF-Processor. The final processed video/GIF is returned in the HTTP response to the mobile application (i.e. keyboard service) in order to be posted to the current social media application. The three Face Swapping approaches (see Sect. 4.4) have been implemented and tested as independent Python modules and one approach was adopted as the optimal one to be enforced by the GIF-processor Python module and integrated in the platform. The Face Swapping algorithm provides the capability to replace the signer’s face with the end-user’s face to enable a more fun and interactive communication in social media.

4.2.1 Content management system

The ASM4Deaf website includes the web platform as the backend, designed specifically for uploading sign language videos and images for face swapping. It is the role of the administrator of the platform to upload a new face image or a new video/GIF or edit the metadata of an existing one. The administrator needs to complete specific input fields on a specific web page. All web pages have a consistent look and feel, as their input fields are integrated into a consistent environment.

An important functionally supported, is the ability to upload a new sign language video. Once the video has been uploaded, the administrator will also need to select the category that best represents the theme of the video. This enables to search for a video/GIF when a user performs a search in the mobile application. Three category examples are: “awesome”, “ThumbsUp” and “wow”. In addition to uploading sign language videos and then assigning them to a category, there is also the need to be able to create a video category in order to collect important metadata about it. This includes the sign language that is being used in the video (e.g. American Sign Language), whether the signer in the video is right/left-handed or using both hands and whether there are subtitles in the video. Such metadata helps in organising the videos in the database, knowing what videos are available in the categories and in which sign language. This enables also the administrator to find the videos, in order to edit them when required.

The web platform also enables the administrator to upload face images of famous people, including those associated with the deaf community and to categorise these. Current categories are “Famous deaf people”, “Sport”, “Movies”, “Science”, “Politics”, “Inspirational people”, “Art”, “Entrepreneurship”, “Humanity”, “Poetry” and “Music”.

Providing the ability to also edit or delete the content uploaded (i.e. videos and images) makes it a simple process for the administrator to search and find specific videos and face

images and provides a platform to organize and manage large amounts of video and face image content. The user-friendly interface makes this it ideal to manage and organize such content.

4.2.2 The web API

The Web API that was created for the purposes of the ASM4Deaf system provides access to a collection of videos and faces and their corresponding information, as well the ability to invoke the face swapping algorithm. The high-level description of the available endpoints and their functionality are described here.⁶ By using these endpoints, developers can retrieve information about videos and face images, either by category or for the entire collection. This information can be used to build a wide range of applications, such as facial recognition systems, video analysis tools, and more. Interested readers can find more details for the API by following the footnote URL.⁷

4.3 Sign language video recordings

A flyer was published to recruit deaf signers for recording the videos and expanding the directory of ASL GIFs. After screening each participant for their signing clarity and language choices, as well as to ensure the signers were of different ethnicities, then they were invited to the small studio. To enhance the diversity of signers, a BIPOC inclusive approach was pursued.

Each participant who officially showed up to the studio was asked to sign a consent form. After the completion of the videos, they were given a stipend.

The GIF style videos were recorded with an iPhone 13 in a small studio with a green screen and lightning. The signers were given a word or a phrase to sign and they would sign it in their own typical fashion. These videos were uploaded to google drive and then to FCP as a set. The files were labeled according to the video file number and the initial of the participant's first name. Then these files and coded directory spreadsheet were sent to the team for the next phase of the process.

4.4 Face swapping

As discussed in [37], three alternative approaches were explored for the creation of the end-result GIF: 1) pre-processing; 2) masking and 3) face swapping. Figures 1, 2 and 3 depict the respective approaches. In pre-processing, a manual video approach is followed where head-torso combinations have already been created and are stored in the system's database. In masking the aim is to help "merge" a GIF with another video or image. In face swapping, only the selected GIF and an image of the face which is to replace that of the signer's is required to produce the final result.

Based on the exploration and experiments done in the context of the project, Table 3 summarises the advantages and disadvantages of each face swapping approach. The

⁶ Web APIs—High-Level Description:

<https://docs.google.com/document/d/17GI2vC6-U4d0PtFZkAHI1tYjRMGsZcZzsJU3Vz9lcpE/edit?usp=sharing>

⁷ API: <https://drive.google.com/file/d/1gTuq9w1TOehYkrtYa0lr3raxHhPM-x00/view?usp=sharing>

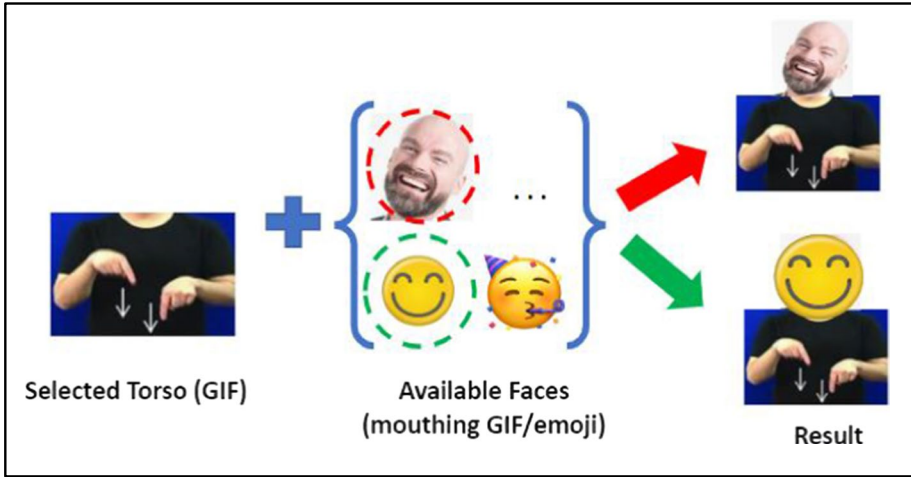


Fig. 1 The pre-processing approach

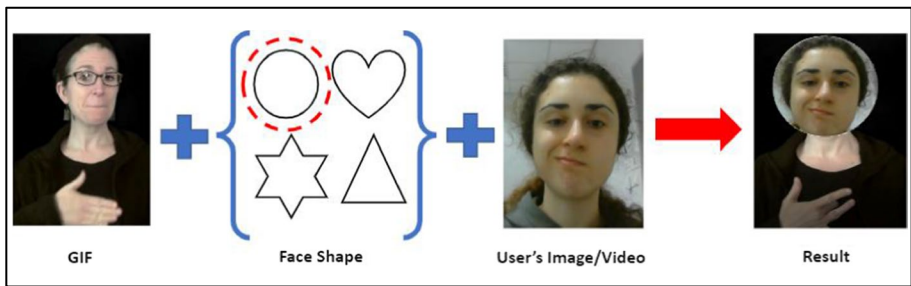


Fig. 2 The masking approach

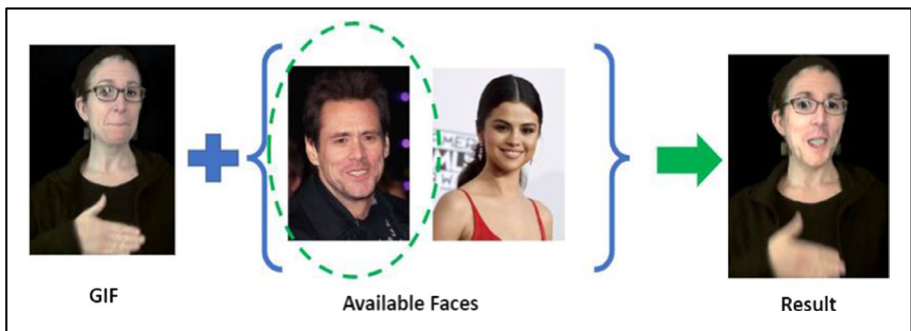


Fig. 3 The face swapping approach

Table 3 Comparison of the three face swapping approaches

Approach	Disadvantages	Advantages
Approach 1 – Pre-processing: Pre-processed GIFs, all possible combinations	<ul style="list-style-type: none"> Needs lots of memory space Needs lots of time for the pre-processing of the GIFs The most expensive approach (pre-processing time, memory space) Difficult to add more torsos and faces in the future Parts of the interpreter's hands are not visible at times (they get hidden behind the face part of the GIF) 	<ul style="list-style-type: none"> Better performance (faster) Better quality
Approach 2—Masking: Result created upon request (face-masking technique)	<ul style="list-style-type: none"> Slower than approach 1 (5–6 s delay to create the video) The result's quality depends on the user's device (i.e. their camera) Parts of the interpreter's hands are not visible at times (they get hidden behind the face part of the GIF) 	<ul style="list-style-type: none"> Needs relatively the least memory space Easy to add new GIF options
Approach 3 – Face swapping: Result created upon request (face swapping technique)	<ul style="list-style-type: none"> The slowest approach (minimum 20–30 s delay to create the video) The result's quality depends on the interpreter's and famous person's face characteristics The result retains the interpreter's facial expression Difficult to implement 	<ul style="list-style-type: none"> Needs relatively little memory space Easy to add new GIF options and famous faces

third approach provides the most viable and sustainable face swapping method since it does not require to manually add these images and does not require so much storage space as in the first approach, while it provides better results when implemented correctly in comparison to the second approach, i.e. parts of the interpreter's hands are not visible at times (they get hidden behind the face part of the GIF).

4.4.1 Image pool

Choosing to proceed with the Face Swapping approach, it was imperative to run several tests to ensure a decent end-result quality. The tests mainly focused on defining a set of necessary (generic) qualifications for the (face) images used with our face swapping tool, and any other (such tool).

Findings from the tests demonstrated: a) a very low level of tolerance in the presence of any material (e.g. accessories, hair) that may conceal any part of the face (forehead included), b) much sensitivity to the image's quality and c) a quite narrow range of acceptable forward-facing diverging head positions.

Also, it was noted that face swapping using pictures of artworks, such as a portrait or a sculpture, often do not produce a satisfying end result and/or completely fail the face swapping procedure. This was quite expected as many artworks can immediately be disqualified according to the a-c criteria mentioned above. It should also be noted, however, that some of them simply failed the procedure at the face landmark recognition phase, since face detecting machine learning models in use are trained to discern mostly real people's faces; in art, the tendency is to find a more abstract way of depicting faces and many materials to choose from to achieve so.

The above qualification criteria were then used to decide upon the images that comprise the application's Image Pool (i.e. the set of images that will be offered to the user). These will also be converted into quality increasing tips provided to the user for when they want to take/upload their own image for the face swapping procedure.

The image collecting procedure included 1) determining the face categories; 2) searching images that can be grouped into each face category; 3) applying the qualification criteria to eliminate images; and 4) deciding whether more images are needed in face category due to elimination process. It should be noted that the aim was to collect 20 images for each face category that had no license restrictions (i.e. using Wikimedia commons). The current face categories are "Famous deaf people", "Sport", "Movies", "Science", "Politics", "Inspirational people", "Art", "Entrepreneurship", "Humanity", "Poetry" and "Music".

A spreadsheet was used in the image collecting procedure. In the main tab of the document, all face categories and respective images collected for those listed. Each face category also had its own tab, where the collected images were saved to be evaluated according to the qualification criteria.

5 HI-FI Prototype

Considering the users' feedback (see Sect. 5) from the Lo-Fi prototype evaluation, a Hi-Fi prototype was then designed and evaluated in turn. The respective screens are presented below.

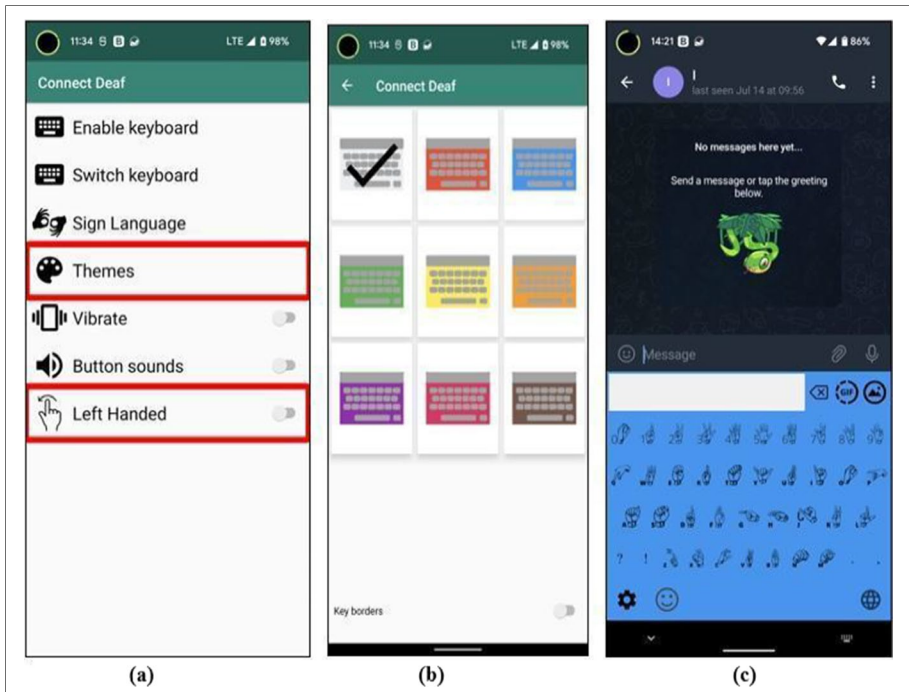


Fig. 4 The main settings screen (a) of the Hi-Fi prototype, theme selection options (b) and view of keyboard with options set as blue theme, right-handed, ASL (c)

In the settings screen (see Fig. 4a), users will first need to enable the keyboard so that it becomes an option for them to select in the MSMAs of preference. Alternatively, they can choose to switch the keyboard from the settings, thus making it the default keyboard of the phone. It is also required to choose the preferred sign language keyboard. The app offers 17 SLA keyboards. In addition, there are different background themes for the keyboards, as can be seen in Fig. 4(a) and (b). This enables them to choose a contrast that is more visually appealing to them, with the foreground being the signs and icons on the keys, and the background the colour behind these. Other options include, adding vibrations for haptic feedback but also sound to buttons for the hearing family members, friends and colleagues of persons who are deaf. Lastly, there is the option to choose left-handed keyboards instead of the default right-handed ones. It should be noted that although the majority of sign language keyboards are one-handed, there are others that use both hands, some of which are also included in the app.

In the scenario that the end-user's MSMA is WhatsApp: By clicking on the keyboard icon at the bottom right-hand corner of the screen, the option will appear to select a keyboard. Once the user selects the Connect Deaf keyboard, the keyboard will change. The user can now type a message using SLA and post it to the chat pane either as an image or GIF. Should the user want to construct a sign language GIF, the emoticon with the smiley face must be selected.

In WhatsApp (see Fig. 5a), the user can search for a GIF with keywords. Connect Deaf app version 2 aims to follow a similar approach in the middle screen (see Fig. 5b),

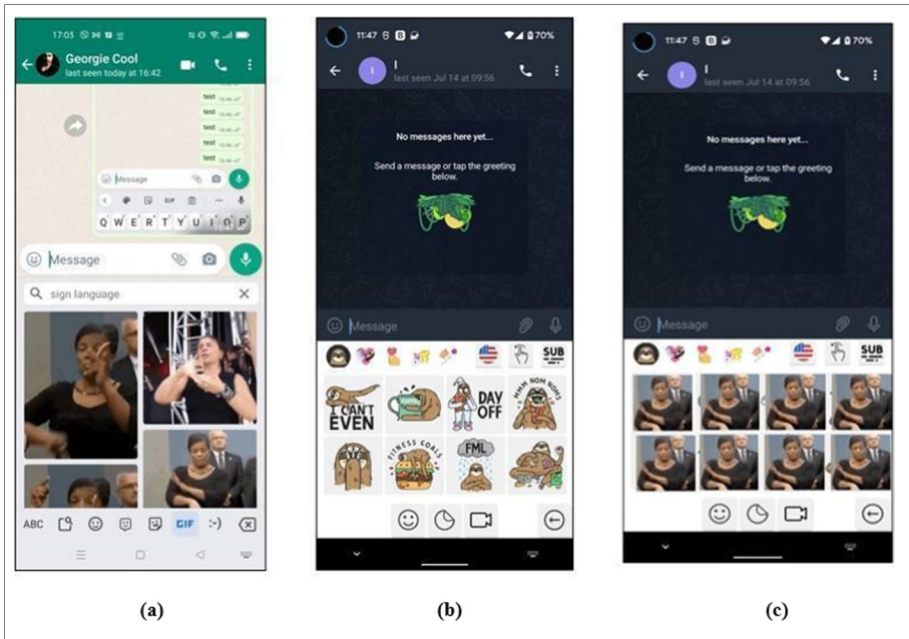


Fig. 5 (a) Searching for a GIF in the Hi-Fi prototype; version with dummy images to evaluate layout of GIFs (b) and version with sign language images to make it more realistic (c)

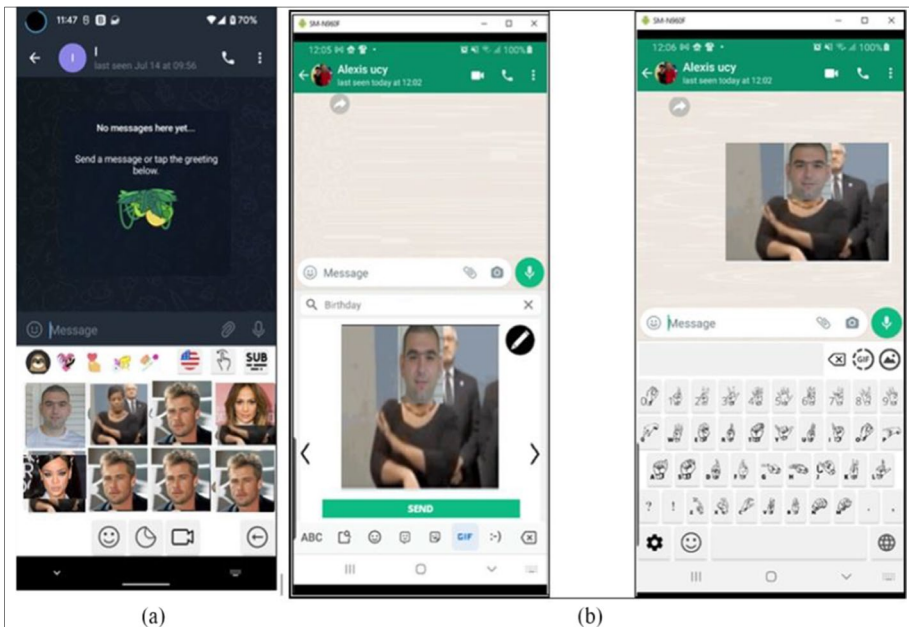


Fig. 6 (a) Selecting a face image to initiate the face swapping process in the Hi-Fi prototype (b) The GIF after the face swapping process that has then been posted to the chat

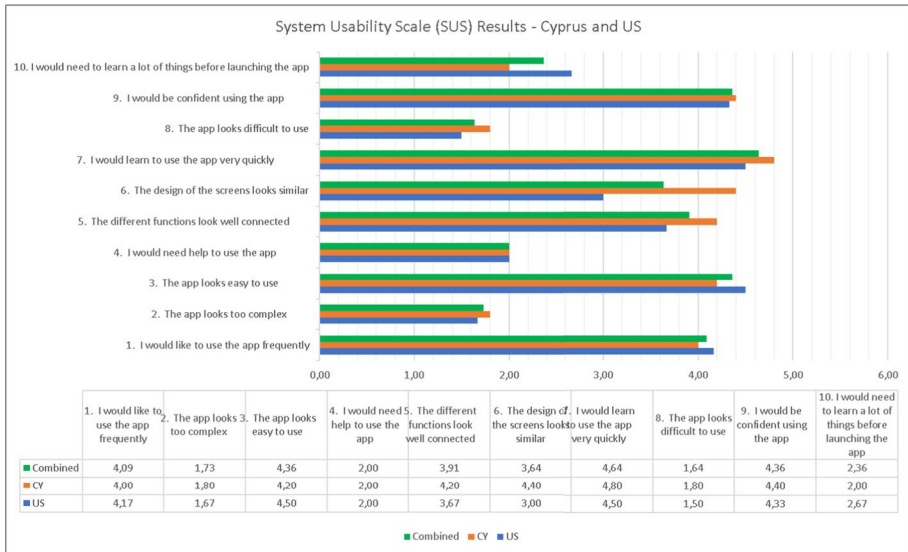


Fig. 7 Comparison of Lo-Fi prototype SUS evaluation by deaf end-users from Cyprus and the US

yet avoids the use of text to type a keyword as part of a search. Instead, icons are used to represent categories of GIFs. There are two reasons for designing it in this way; first, to avoid text use, since sign language is the primary language of the main users, and secondly, from a more technical approach, the number of videos (about 1000+) at this stage of development are relatively a limited number. Thus, to avoid user frustration, by typing keywords that may not yield any results, it was preferred to create categories that will contain videos. Once the user selects an icon (i.e. category) from the top menu bar, the respective GIFs of that category are presented in tiles. An example, with a sign language user, is presented on the right screen (see Fig. 5c).

Once the GIF has been selected by the user in screen 2, they will proceed to screen 3 (see Fig. 6a), which is a similar screen to that of screen 2 in terms of design, but instead of presenting the sign language GIFs, it presents images of famous peoples' faces. The user can on this screen choose again a category, i.e. sportsman/sportsman, actor/actress, famous deaf person, etc. from which the face will be chosen for face swapping, i.e. to swap the face of the signer from the original video with that of the famous person. There is also the option for users to have an image of their own face, which could also be used for face swapping.

If the user is satisfied with the face swapping result, it can then be posted to the chat pane of the MSMA that is being used. Otherwise, the GIF could be edited (see Fig. 6b).

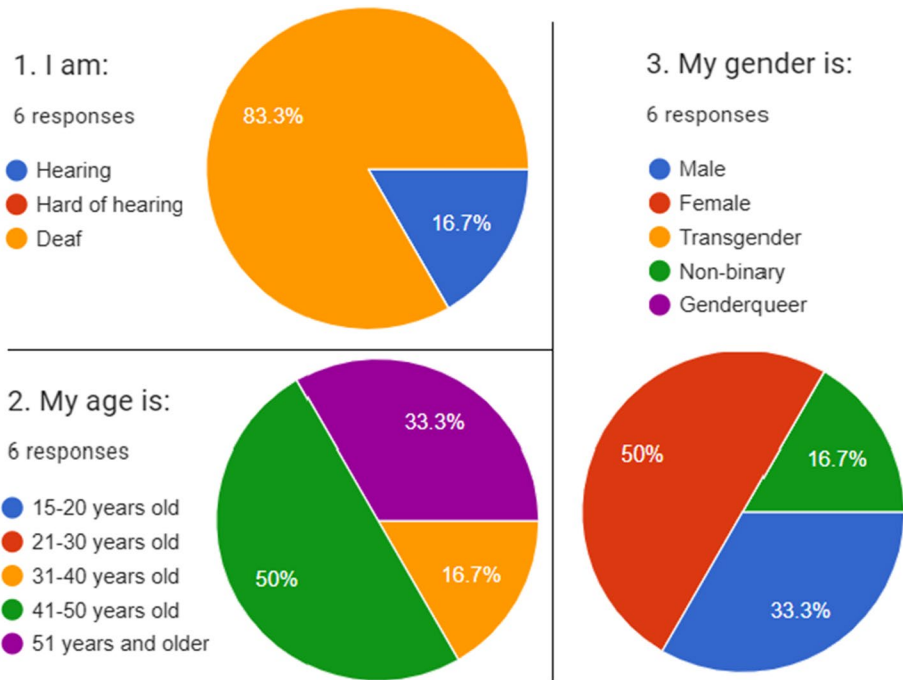


Fig. 8 Demographics of participants in US focus group

6 Results

6.1 Low-fidelity prototypes

This section presents to the reader the procedure and the results of the focus groups that were performed in Cyprus and US for the evaluation of the Lo-Fi prototypes. The following graph showcases the results from the Cyprus, US and combined results from the Lo-Fi prototypes focus groups.

Figure 7 showcases that positive items of the SUS instrument (the odd items) received scores from participants (on average) well above 4 out of 5, with the only exception being item 5 in which the results in CY are well above (4,20/5) the results in the US (3,67/5) that can potentially be an indicator that the transition from one screen to the next needed to be improved prior to application implementation. This result was positively received and the different screens were evaluated from the project partners in an effort to improve the usability of the mobile application. Moreover, the graph above showcases that negative items of the SUS instrument (the even items) received scores from participants (on average) close to 2 out of 5, with the only exception being item 6 in which the results are relatively high. The issue with item 6 is that the statement is confusing for the participants and while the intention was to evaluate if the users find inconsistency between the different screens, the participants interpreted the statement as asking exactly the opposite.

Nevertheless, the overall SUS scores in CY, US and the combined scores were calculated as follows and showcase that the Lo-Fi prototypes were assessed with Grade B and

1. I would like to use the app frequently

6 responses

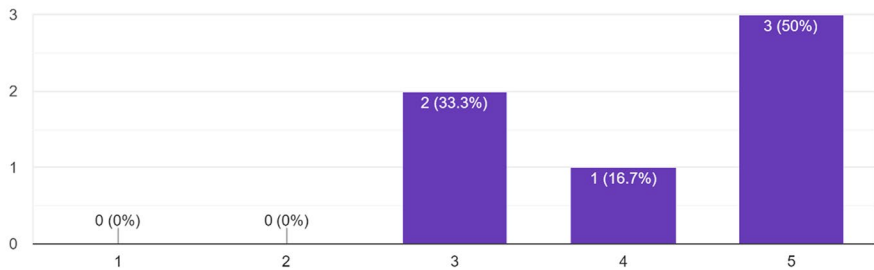


Fig. 9 Application use responses

Adjective Rating of Good in accordance to the aforementioned table that serves as a general guideline on the interpretation of SUS score.

$$SUS\ Score(CY) = 74.00, \quad SUS\ Score(US) = 75.83, \quad SUS\ Score(Combined) = 75.00$$

6.1.1 US focus group

In this subsection only the US Low-Fidelity focus group and its results are described, as the CY Low-Fidelity focus group and its results are described in the related conference paper [37]. The US focus group was organized in collaboration with AnnRae Consulting (ARC) LLC, a consultant agency for deaf people based in the United States. The focus group was planned to take place online, since the aim was to attract more people. A total of 12 participants had registered for the focus group, with 6 joining on the day. From the 6 participants, 5 participants represented the primary end-user group and 1 the secondary end-user group. All participants completed the adapted SUS evaluation instrument.

The demographic data of the people that answered the questionnaire are shown in Fig. 8. The responses are 83.3% from people who are deaf or Hearing Impaired (i.e. primary end-user group) and 16.7% from Hearing people (i.e. secondary end-user group). It is also important that we have coverage on all age groups, as well as the fact that all participants have a good and very good use of mobile applications and smartphones, i.e. the main target group.

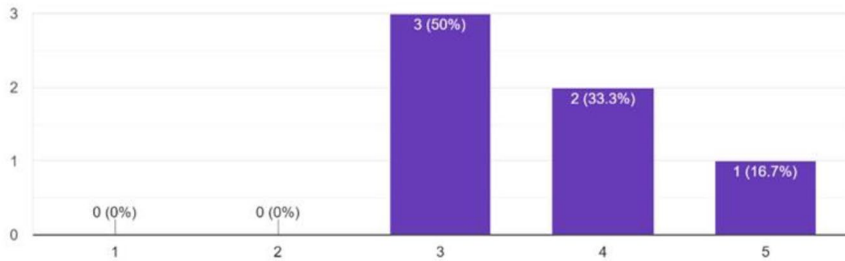
Despite the small number of responses, participants are confident that they would like to use the application as exhibited in Fig. 9. In fact, 66.7% of the participants replied that they would like to use the application often and very often. This showcases a real need for such an application that enables deaf, hearing impaired and hearing end-users to communicate socially in their preferred sign language through MSMAs.

Two more important results from the questionnaire are that the Lo-Fi prototypes were assessed as well thought and well connected, as well as the fact that the users will have no problems in using the mobile application (see Fig. 10). This revealed that the design of the Lo-Fi prototypes was intuitive and well understood in terms of the functionality to be offered by the mobile application, by all participants.

Despite the fact that the responses were received from a small number of participants ($N=6$), still the above results showcase a strong interest for the Connect Deaf mobile app.

5. The different functions look well connected

6 responses



7. I would learn to use the app very quickly

6 responses

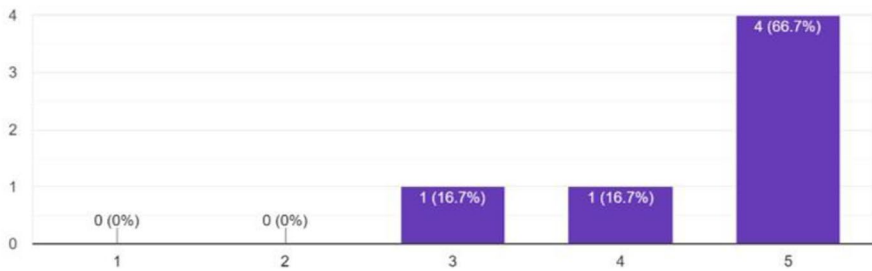


Fig. 10 Usability and Ease of Use responses

6.2 High-fidelity prototypes – focus groups in CY and US

This section presents to the reader the procedure and the results of the focus groups that were performed in Cyprus and US for the evaluation of the Hi-Fi prototype. The results are presented together since in Cyprus 6 participants joined the Hi-Fi focus group but only one participant has answered the adapted SUS instrument following the focus group. In the US Hi-Fi focus group 12 participants joined and 5 participants have answered the adapted SUS instrument.

In terms of demographic data from the US, 4 of the participants were from the primary end-user group (i.e. deaf or Hard Hearing) and 1 of the participants was from the secondary end-user group (i.e. Hearing). The participant that answered the questionnaire from CY was also from the primary group. The representation was from various age groups as can be seen in Fig. 11.

The participants had an excellent knowledge and use of smartphone and mobile applications that makes the results of the focus group very useful, irrespective of the small sample size. The same applies for the CY participant. The results of the Hi-Fi prototypes evaluation focus groups in terms of usability were recorded with a SUS score of: SUS Score (combined)=67. Despite the lower score than the Lo-Fi focus groups the graphs below (i.e., Figs. 12–13) indicate very positive feedback and the fact that the consortium team has worked and incorporated feedback from the Lo-Fi prototype evaluation focus groups. In fact, 60% of the participants indicated that they are very likely to use the app, while 80% were very confident it would be easy to use and for learning easily to use the mobile application.

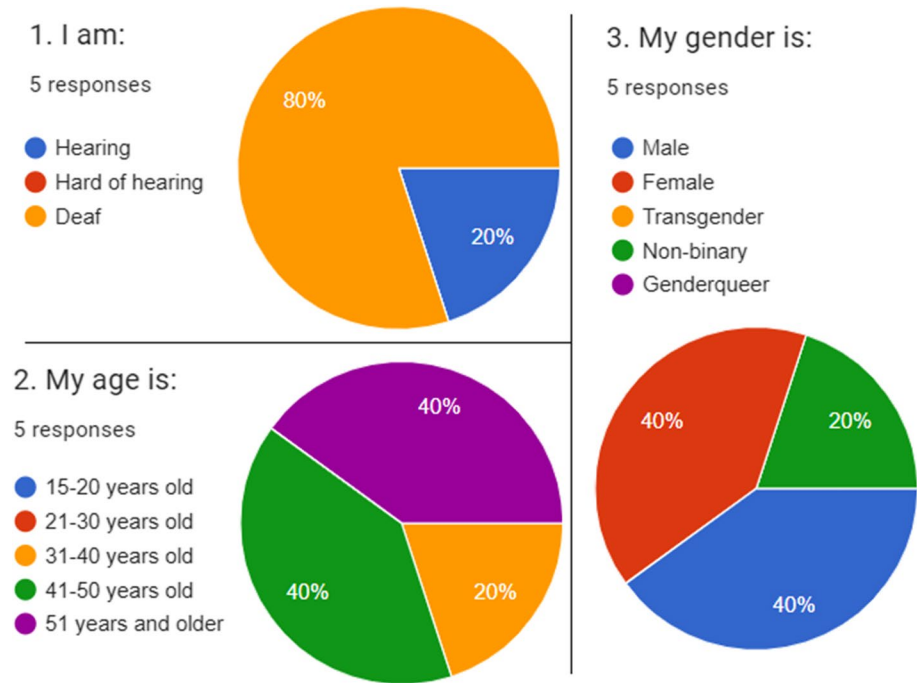


Fig. 11 Demographics of participants

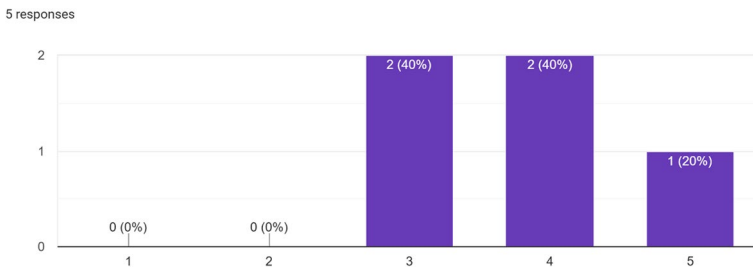


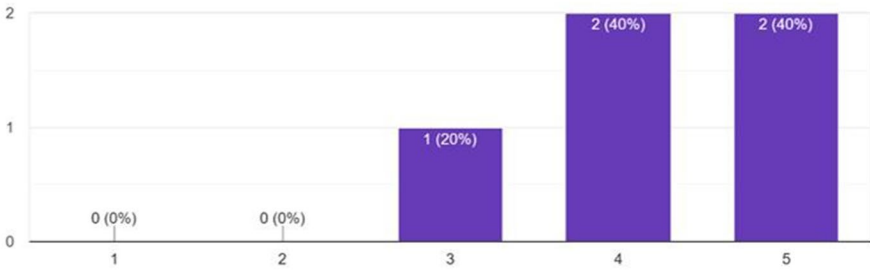
Fig. 12 Application use responses

6.3 Functional prototype

Based on the evaluation of the Lo-Fi and Hi-Fi prototypes and the results, the final mobile application was designed and developed for the Android and iOS platforms. The mobile application is installed as a keyboard in the Android and iOS platforms and can be used in several social media, e.g. WhatsApp, Telegram, Messenger, Google apps.

3. The app looks easy to use

5 responses



7. I would learn to use the app very quickly

5 responses

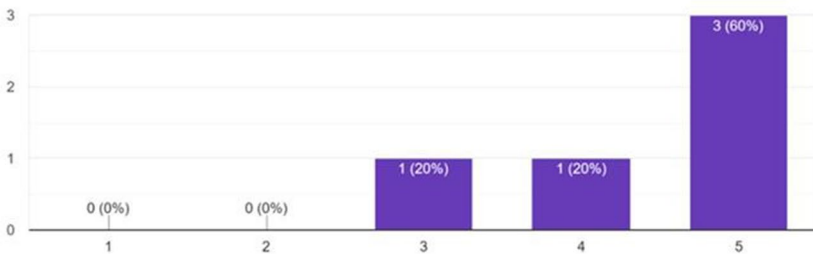


Fig. 13 Usability and Ease of Use responses



Fig. 14 Mobile application SLA, emoticon and sticker features

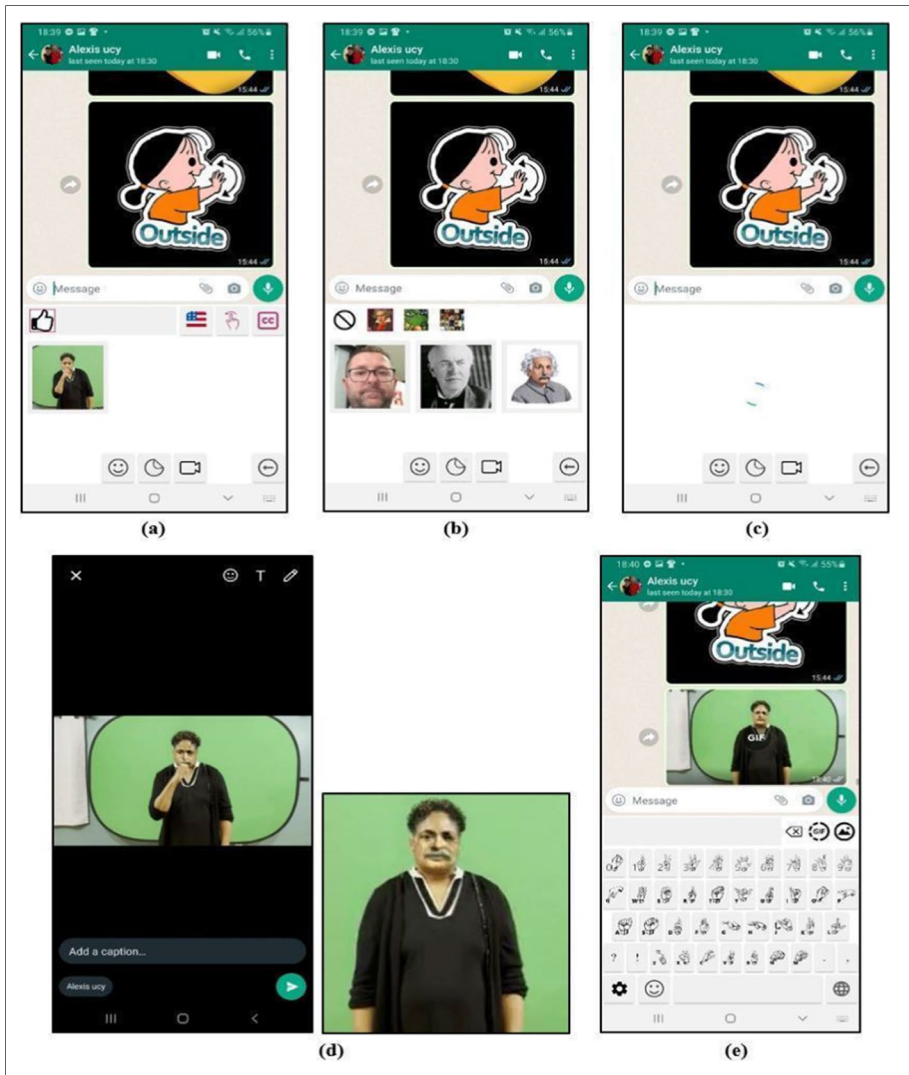


Fig. 15 Face swap video-based sign language feature

The keyboard is enabled by clicking the keyboard button at the bottom right of Fig. 14 (a) and selecting the Connect Deaf keyboard. The SLA Keyboard enables to type a message using the SLA, and the user can then click on either the GIF or Image buttons (next to the backspace button) shown on Fig. 14 (a) to respectively generate a GIF or Image representation that can be posted in social media (see Fig. 14a). The keyboard enables the user to select and send emoticons (see Fig. 14b) and select and send stickers (see Fig. 14c).

The mobile application provides an additional very important feature that enables users to communicate in a fun and intuitive way in social media. By clicking on the video icon/button of the application (see Fig. 15 (a)), the user is presented with the screen shown that allows selecting the language of the video, enable or disable if the signer is left handed in the video and finally enable or disable if the video has subtitles. Based on this filtering

using the three icons/buttons on the top the available videos are shown. The user then selects the video and moves on to the next screen (see Fig. 15b) where the user has the possibility to click on one of the top face categories (e.g. celebrities) to select one face or even take a custom selfie to use it for the face swapping. Once the face is selected, in this case the Einstein face was selected, the user sees the next processing screen (see Fig. 15c) and as soon as the face swapping is completed (i.e. Einstein's face swapped with that of the signer in the original video) then the user is presented with the final video result (see Fig. 15d) that the user can edit, delete or approve and send to be posted on social media (see Fig. 15e). The mobile application is currently functional, the videos recorded in the ASL language are currently being segmented and uploaded to the application and the final user testing will begin.







During internal testing, it was determined that the face swapping processing required on average 1 min:40 s and this caused timeout errors in many cases. Moreover, while the average video added via the platform (i.e. CMS for administrator) was about 1 MB in size initially, the video returned as a result of the face swapping was almost 50 MB. When it was received and sent on social media (e.g. WhatsApp), it was compressed to ~250 kb by the social media app. This issue, i.e. the video size returned as a result of the face swapping being large in size, required to compress the video. Two options to address this size issue were: to implement configurations in the face swapping tool/code or use a library. After resolving the issue, the GIF size was measuring between 2 and 5 MB in size, depending on the duration of the video. The time needed to perform face swapping has also been reduced from approximately 1 min:30 s to 30 s, depending on the duration of the video. As part of future work, it was determined that the duration could be further reduced by using multi-threading for the face swapping. We conclude with highlighting the main features of the Connect Deaf app version 2, that forms part of the ASM4Deaf system in Table 4.

7 Conclusion and future work

The work in this paper presents the ASM4Deaf system: Connect Deaf mobile app version 2, content management system and Web API, that are evaluated in the context of the focus groups. The main contribution of the paper refers to the design and development of the Connect Deaf mobile app (i.e. as a keyboard service) that provides accessibility and enables the social inclusion of deaf end-users in popular social media applications such as WhatsApp, Viber, Messenger and Telegram. The results of the evaluation reveal that, in fact, 60% of the participants are very likely to use the app, 80% were very confident that the app is easy to use and that they can easily learn how to use the mobile application. Moreover the evaluation of the prototypes resulted in a SUS Score (combined) of 67. Despite the low number of participants, the results are important since the people had an excellent knowledge and use of smartphone and mobile applications.

The aforesaid results were achieved through the central and important co-creation HCD approach that was followed in this work in order to take fully into consideration the requirements of the deaf end-users. This led to: 1) design and development of the ASM4Deaf system, and its Connect Deaf mobile app version 2, forming part of the overall system; enabling use of SLA keyboards in MSMA in 17 different sign languages and 2) the evaluations of Low- and High-Fidelity (Lo/Hi-Fi) prototypes aimed to enhance the app's design and functionality, i.e. the ability to browse, search and edit animated (i.e. not static)

Table 4 The app’s main features

Feature	Example
<p>Post ASL words and phrases as GIFs in MSMAs.</p>	
<p>Swap the face of the signer of an ASL GIFs with the end-user’s face or that of a famous person, including famous deaf people, sportsmen and sportswomen, actors, politicians, poets and singers while keeping the facial expressions of the originally signed GIF.</p>	
<p>Create and post 2D SLA messages as GIFs and images in 17 different SLAs.</p>	
<p>Create and post 3D messages as GIFs and images in 2 different SLAs: American and South African.</p>	
<p>Use it in many favourite MSMAs: including Telegram, WhatsApp, Google Hangouts, Instagram, Viber, Facebook Messenger, Facebook, Skype, MMS.</p>	
<p>Send sign language stickers and emoticons to personal networks (deaf and hearing).</p>	

GIFs in American Sign Language (ASL), and use face swapping. This makes the full set of features offered by social media apps accessible to deaf end-users and their personal network (i.e. friends and family), enabling them to utilise all communication capabilities in social media apps, which is the main goal and contribution of this work.

Limitations of this study include the small number of participants in the focus groups in comparison to the number registered, and the small sample size who had answered the adapted SUS questionnaire. Nevertheless, the participants were actively engaged and contributed with insightful comments and provided valuable input in the co-creation process. It should also be noted that despite the small number of participants, each focus group session organized was still in line with Nielsen's [39] suggestions; that it is more effective when the group consists of six to nine users and that the session lasts for about two hours.

Finally, future work includes the following extensions to the ASM4Deaf system. The first concept is to extend this research work by providing the ability to the end-users, by following a crowd-sourcing approach, to contribute their own sign language words and phrases to the system, so that they can be shared and used by other deaf end-users with the help of the face swapping algorithm proposed in this work. A second concept is to proceed with research work that enables to use artificial intelligence and machine learning in order to automatically or semi-automatically transform typed text or SLA input on the keyboard to sign language videos (e.g. using sign language avatars or specific human sign language interpreters) for educational (i.e. family and friends) and social purposes (i.e. deaf end-users). Extending the number of keyboards, words and phrases in ASL and other sign languages are also future work perspectives.

Acknowledgements The project PRE-SEED/0719(B)/0278 is co-financed by the European Regional Development Fund and the Republic of Cyprus through the Research and Innovation Foundation.

Funding Open access funding provided by the Cyprus Libraries Consortium (CLC).

Data availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. World Health Organization (2023) Deafness and hearing loss. Retrieved October 9, 2023 from <https://www.who.int/en/news-room/fact-sheets/detail/deafness-and-hearing-loss>
2. World Federation of the Deaf (n.d.) Deafness and hearing loss. Retrieved October 9, 2023 from <https://wfdeaf.org/who-we-are/>
3. European Union of the Deaf (n.d.) Members. Retrieved October 9, 2023 from <https://www.eud.eu/eud/full-members/>
4. Mitchell R, Young T, Bachleda B, Karchmer M (2006) How many people use ASL in the United States? Why estimates need updating. *Sign Lang Stud* 6(3):306–333. <http://www.jstor.org/stable/26190621>

5. Debevc M, Kožuh I, Kosec P, Rotovnik M, Holzinger A (2012) Sign language multimedia based interaction for aurally handicapped people. In: *Computers Helping People with Special Needs* (pp 213–220). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-31534-3_33
6. Yeratziotis A, Achilleos A, Koumou S, Thibodeau RA, Vanezi E, Geratziotis G, Papadopoulos GA (2022) Accessible system and social media mobile application for deaf users: ASM4Deaf. In: *Proceedings of the 2022 ACM Conference on Information Technology for Social Good (GoodIT '22)*. Association for Computing Machinery, New York, pp 39–47. <https://doi.org/10.1145/3524458.3547234>
7. Yeratziotis G, Van Greunen D (2013) Making ICT accessible for the deaf. In: *2013 IST-Africa conference & exhibition, Nairobi*, pp 1–9
8. Yeratziotis A, Zaphiris P (2018) A heuristic evaluation for deaf web user experience (HE4DWUX). *Int J Hum-Comput Interact* 34(3):195–217
9. Kemelmacher-Shlizerman I (2016) Transfiguring portraits. *ACM Trans Graph (TOG)* 35(4):1–8
10. Yeratziotis A, Zaphiris P (2015) Interactive software technology for deaf users: mapping the hci research landscape that focuses on accessibility. In: Antona M, Stephanidis C (eds) *Universal access in human-computer interaction. Access to today's technologies. UAHCI 2015. Lecture Notes in Computer Science()*, vol 9175. Springer, Cham. https://doi.org/10.1007/978-3-319-20678-3_25
11. Pieri K, Cobb SVG (2019) Mobile app communication aid for Cypriot deaf people. *J Enabling Technol* 13(2):70–81. <https://doi.org/10.1108/JET-12-2018-0058>
12. Cabanillas-Carbonell M, Cusi-Ruiz P, Prudencio-Galvez D, Herrera Salazar JL (2022) Mobile Application with Augmented Reality to Improve the Process of Learning Sign Language. *Int J Interact Mob Technol (IJIM)* 16(11):51–64. <https://doi.org/10.3991/ijim.v16i11.29717>
13. Samonte MJC, Gazmin RA, Soriano JDS, Valencia MNO (2019) BridgeApp: An assistive mobile communication application for the deaf and mute. In: *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, pp 1310–1315. <https://doi.org/10.1109/ICTC46691.2019.8939866>
14. Web Accessibility Initiative (2012) *Web Content Accessibility Guidelines (WCAG)*. Retrieved October 9, 2023 from <https://www.w3.org/WAI/>
15. Efthimiou E, Fontinea S-E, Hanke T, Glauert J, Bowden R, Braffort A, Collet C, Maragos P, Goudenove F (2010) Dicta-sign–sign language recognition, generation and modelling: a research effort with applications in deaf communication. In: *Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, pp 80–83
16. Efthimiou E et al (2009) Sign Language Recognition, Generation, and Modelling: A Research Effort with Applications in Deaf Communication. In: Stephanidis C (ed) *Universal access in human-computer interaction. Addressing diversity. UAHCI 2009, Lecture notes in computer science*, vol 5614. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02707-9_3
17. Elliott R, Glauert JRW, Kennaway JR, Marshall I (2000) The development of language processing support for the ViSiCAST project. In: 108 (ed) *Proceedings of the fourth international ACM conference on Assistive technologies (Assets '00)*. Association for Computing Machinery, New York, p 101. <https://doi.org/10.1145/354324.354349>
18. Elliott R, Glauert JR, Kennaway JR, Marshall I, Safar E (2008) Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. *Univ Access Inf Soc* 6(4):375–391
19. Fels D, Gerdzhev M, Hibbard E, Goodrum A, Richards J, Hardman J, Thompson N (2009) Sign language online with Signlink Studio 2.0. *HCII LNCS* 5616:492–501
20. Monteiro CDD, Gutierrez-Osuna R, Shipman FM (2012) Design and evaluation of classifier for identifying sign language videos in video sharing sites. In: *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12)*. Association for computing machinery, New York, pp 191–198. <https://doi.org/10.1145/2384916.2384950>
21. Allen TE, Letteri A, Choi SH, Dang D (2014) Early visual language exposure and emergent literacy in preschool deaf children: Findings from a national longitudinal study. *Am Ann Deaf* 159(4):346–358
22. Neidle C, Sclaroff S, Athitsos V (2001) SignStream™: A Tool for Linguistic and Computer Vision Research on Visual-Gestural Language Data. In: *Behavior Research Methods, Instruments, and Computers* 33(3):311–320
23. Bottoni P, Borgia F, Buccarella D et al (2013) Stories and signs in an e-learning environment for deaf people. *Univ Access Inf Soc* 12:369–386. <https://doi.org/10.1007/s10209-012-0283-y>

24. van Zijl L (2006) South African sign language machine translation project. In: Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility (Assets '06). Association for Computing Machinery, New York, pp 233–234. <https://doi.org/10.1145/1168987.1169031>
25. Barberis D, Garazzino N, Prinetto P, Tiotto G (2011) Improving accessibility for deaf people: an editor for computer assisted translation through virtual avatars. In: The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '11). Association for computing machinery, New York, pp 253–254. <https://doi.org/10.1145/2049536.2049593>
26. Alexander O, Rogers M, Lambeth W, Chiang M, Debevec P (2009) Creating a photoreal digital actor: the digital Emily project, Conference for visual media production, vol 2009, London, pp 176–187. <https://doi.org/10.1109/CVMP.2009.29>
27. Reface.ai, Reface: face-swap app & AI avatar generator. Retrieved October 9, 2023 from <https://hey.reface.ai/>
28. Mosaddegh S, Simon L, Jurie F (2015) Photorealistic face de-identification by aggregating donors' face components. In Computer Vision--ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1–5, 2014, Revised Selected Papers, Part III 12 (pp 159–174). Springer International Publishing
29. Naruniec J, Helminger L, Schroers C, Weber RM (2020) High-resolution neural face swapping for visual effects. *Comput Graph Forum* 39:173–184. <https://doi.org/10.1111/cgf.14062>
30. Zhu Y, Li Q, Wang J, Xu CZ, Sun Z (2021) One shot face swapping on megapixels. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp 4834–4844)
31. Nguyen TT et al (2022) Deep learning for deepfakes creation and detection: A survey, computer vision and image understanding. *Comput Vis Image Underst* 223(103525):1077–3142. <https://doi.org/10.1016/j.cviu.2022.103525>
32. Chesakov D, Maltseva A, Groshev A, Kuznetsov A, Dimitrov D (2022) A new face swap method for image and video domains: a technical report. arXiv preprint [arXiv:2202.03046](https://arxiv.org/abs/2202.03046)
33. Chen D, Chen Q, Wu J, Yu X, Tong J (2019) Face swapping: realistic image synthesis based on facial landmarks alignment. *Math Probl Eng*
34. Groshev A, Maltseva A, Chesakov D, Kuznetsov A, Dimitrov D (2022) GHOST—a new face swap approach for image and video domains. *IEEE Access* 10:83452–83462. <https://doi.org/10.1109/ACCESS.2022.3196668>
35. Ding X, Raziei Z, Larson EC, Olinick EV, Krueger P, Hahsler M (2020) Swapped face detection using deep learning and subjective assessment. *EURASIP J Inf Secur* 2020(1):1–12
36. Hassani A, Malik H, Diedrich J (2022) Efficiently Mitigating Face-Swap-Attacks: Compressed-PRNU Verification with Sub-Zones. *Technologies* 10(2):46
37. Bijl-Brouwer MVD, Dorst K (2017) Advancing the strategic impact of human-centred design. *Des Stud* 53:1–23. <https://doi.org/10.1016/j.destud.2017.06.003>
38. ISO 9241–210 (2019) Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems
39. Nielsen J (2020) Focus Groups in UX Research. Retrieved October 9, 2023 from <https://www.nngroup.com/articles/focus-groups>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Alexandros Yeratziotis¹  · Achilleas Achilleos² · Stavroulla Koumou³ · George Zampas³ · Regan A. Thibodeau⁴ · George Geratziotis¹ · George A. Papadopoulos³ · Iasonos Iasonas² · Christophoros Kronis²

✉ Alexandros Yeratziotis
alexis@connectdeaf.com

Achilleas Achilleos
achilleas.achilleos@frederick.ac.cy

Stavroulla Koumou
skoumo01@ucy.ac.cy

George Zampas
gzampa01@ucy.ac.cy

Regan A. Thibodeau
arlingui@gmail.com

George Geratziotis
george@connectdeaf.com

George A. Papadopoulos
george@ucy.ac.cy

Iasonos Iasonas
com.ii@frederick.ac.cy

Christophoros Kronis
c.kronis@frederick.ac.cy

¹ A.G. Connect Deaf Limited, 8 Lakonias Street, 2021 Nicosia, Cyprus

² Department of Electrical Engineering, Computer Engineering and Informatics, Frederick University, 7, Y. Frederickou Str. Pallouriotisa, 1036 Nicosia, Cyprus

³ Department of Computer Science, University of Cyprus, 2109 Nicosia, Cyprus

⁴ Annrae Consulting, LLC., PO Box 352, Westbrook, ME 04062, USA