

Innovative Multimodal Translation: Unveiling the Figure Out Application's Real-Time Language Solution

Paula Escudeiro¹, Márcia Campos Gouveia¹, Francisca Escudeiro²,
and Nuno Escudeiro¹

¹GILT – Games, Interaction & Learning Technologies, ISEP, IPP, Porto, Portugal

²ICBAS, Instituto de Ciências Biomédicas Abel Salazar da Universidade do Porto,
Portugal

ABSTRACT

Communication across languages, particularly between deaf and hearing individuals, presents significant challenges. Traditional translation tools, while helpful, often fail to fully address the diverse needs of these users. This paper introduces Figure Out, a mobile application designed to provide real-time translations of text captured from images. By utilizing advanced optical character recognition (OCR), Figure Out translates written text into multiple formats—audio, text, and sign language—making it particularly accessible for individuals with hearing impairments. The app's key advantage is its support for both spoken and signed languages, offering a comprehensive tool for inclusive communication. By integrating text, spoken, and signed language translations, Figure Out takes a unique approach to breaking down communication barriers and fostering inclusion.

Keywords: Communication technology, Inclusive technology, Automatic translation, Accessibility, Sign language

INTRODUCTION

Effective communication in languages other than one's native tongue is often a challenging attempt. Even with a basic grasp of a foreign language, maintaining proficiency can be difficult unless one is a frequent user. Vocabulary and grammar rules tend to fade over time, further complicating communication. This challenge becomes even more pronounced when individuals communicate across distinct languages, especially in cases where the communication involves deaf and non-deaf individuals (Zheng et al., 2020). Translating between spoken and signed languages is complex due to different grammatical rules. Robust systems are needed to facilitate communication between hearing and Deaf individuals, translating spoken languages into sign languages (Stoll et al., 2019). However, we anticipate that technology can play a pivotal role in addressing these challenges (House, 2016), and one such solution is the Figure Out application.

Figure Out is a mobile application designed to facilitate language translation. It accomplishes this by automatically capturing text through the mobile camera and providing translations to the desired language.

What sets Figure Out apart is its unique capability to translate not only between spoken languages but also to and from sign languages. This application has the potential to empower everyday users, including tourists, students, and individuals from all walks of life, by granting them access to information in their preferred language, fostering inclusion, accessibility, and communication.

METHODOLOGY

Figure Out's aim is to enhance accessibility to cultural heritage, offering digital tools and content in various languages, including sign language, for online and mobile access (Escudeiro et al., 2022; Escudeiro et al., 2022). The implementation of an augmented reality application for promoting cultural heritage in sign language was guided by a specific methodology. Additionally, Figure Out provides an online collaborative platform that facilitates the entire tourist experience, including promotion, awareness-raising, site visits, and feedback collection.

Concept

The core of the Figure Out project is a mobile application designed to deliver real-time translations of small text segments captured as images within an augmented reality context. Users can point their mobile device's camera at the text, and the app swiftly translates it into their chosen language. What sets Figure Out apart is its focus on sign language, offering translations as written text and audio to cater to diverse user preferences. The app accommodates various sign languages, enhancing communication inclusivity.

Beyond translation, Figure Out provides a mobile diary tool, allowing users to document thoughts, actions, and surroundings via text, spoken messages, pictures, or videos. Each entry includes transcriptions, audio files, timestamps, geo-tags, participant IDs, and message durations, contributing valuable data for research.

The ultimate goal is to implement these tools within a Living Lab, an open innovation environment where users actively participate in the innovation process. This user-centric approach ensures the technology meets real-world needs effectively.

Figure Out's innovation lies in its real-time sign language translation, distinguishing it from other tools primarily focusing on text or spoken language. Its application in scenarios like reading signs, timetables, tickets, receipts, and menus makes it indispensable for those relying on sign language. Notably, Figure Out builds upon VirtualSign technology, converting written and spoken language into sign language, with real-time translation for seamless communication. VirtualSign has significantly impacted education, providing access to materials for deaf students and promoting inclusivity in various aspects of life.

Figure Out, as an extension of VirtualSign, integrates real-time image-based translation and mobile diary functionality, offering a comprehensive solution for the deaf community (Oliveira et al., 2019a; Oliveira et al., 2019b; Escudeiro et al., 2013).

Objectives

The primary objective of the Figure Out application is to simplify information access, which in turn enhances accessibility, inclusivity, and communication, especially among the deaf and non-deaf communities. With Figure Out, users gain the ability to access information in their first language or a language of their choice. This can have transformative effects on various aspects of society. In cultural settings such as museums and art galleries, it acts as a bridge, enabling deaf individuals to comprehend displayed information and ensuring inclusive and enriching cultural experiences. For tourists and international students, Figure Out functions as a language companion, breaking down language barriers and enhancing the overall experience of international mobility. Most significantly, Figure Out plays a crucial role in reducing communication barriers between the deaf and non-deaf communities, fostering greater accessibility, and understanding.

To better achieve it, the Figure Out Project encompasses the planning, design, construction, and implementation of a comprehensive system that combines a mobile application and website to facilitate the translation of written words into a selected language. This project aims to provide users with translated content in various formats, including written, audio, and sign language. Moreover, the mobile application records and associates the translated words with GPS location data, which is subsequently stored in a central database for analytical purposes. The website complements the application by offering users access to usage statistics, including details on frequently translated words and geographical regions where the application sees the most use.

With a multifaceted approach, Figure Out addresses various user needs and preferences, leveraging real-time data collection and analytical capabilities. By providing multilingual support and enabling easy data translation, it aims to enhance accessibility, user engagement, and the overall user experience in both mobile and web environments.

TECHNICAL RESOURCES

Acknowledging the technical knowledge necessary for mobile development, the selection of a framework that could efficiently serve both platforms with substantial code reusability was deemed appropriate. Following an analysis of the available platforms, the decision was made to opt for PhoneGap.

Frontend Technology

In fulfilling front-end requirements for software development, two key aspects are typically considered: application and website development. For the application, we chose Ionic, an open-source SDK tailored for hybrid mobile app development, ensuring compatibility across Android and iOS platforms. This approach allows us to reach a broader audience without investing in separate platform-specific efforts. Ionic's advantage lies in codebase reuse, enabling efficient development with a consistent user interface. It also supports desktop website creation, leveraging the same technology for versatile applications.

For website development, Angular was selected due to its similarities with Ionic, expediting the process for our experienced development team. Angular, an open-source framework for dynamic web applications, is based on TypeScript, enhancing JavaScript with static typing and advanced features. This choice streamlines codebase writing and maintenance, providing superior tooling and refactoring capabilities. Angular empowers us to create responsive, interactive websites adaptable to diverse screen sizes and user inputs, ensuring performance and scalability.

Backend Server and Technology

When it comes to the backend requirements of our project, our team has decided to implement a NodeJS server. This particular technology is widely used and has a plethora of documentation available, making it easy for our team to work with. By choosing NodeJS, we have been able to create a centralized system that provides more control and easier maintenance of our REST API routes.

To store our data, our team has opted for firebase. This decision was based on the fact that some members of our development team already had experience using this technology. In addition, we found that firebase is highly reliable and provides a secure environment for our data.

System Architecture

Figure Out was initially envisioned as a centralized web service that provides constant access to data and resources across three logical layers: UI, Server, and Data. The UI layer interacts with users through the mobile app or website, while the web service layer handles UI requests and translation. The data layer serves as a persistence mechanism for statistical purposes.

The technology selection was based on the development team's skills, with Ionic chosen for the mobile app due to its cross-platform support, Node.js and Express for the backend server, and Firebase for the database, enabling push notifications and user management. For the sign language viewpoint, the VirtualSign API enables conversion of text to sign language as seen in the following figure.

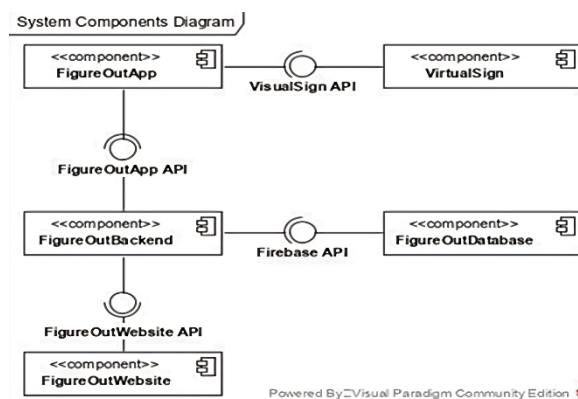


Figure 1: Logical architecture.

The Figure Out mobile application prioritizes user-friendliness and intuitiveness, with carefully crafted prototype mock-ups ensuring a seamless experience. The application's workflow is straightforward: users capture a photo using their mobile camera, configure the translation if needed, and proceed. The app's homepage grants direct access to the translation feature for added convenience.

Beyond translation, users can explore the About page for additional information about Figure Out and its features, enhancing their understanding of the app's capabilities. When initiating a translation, users capture a photo of the text, and the app identifies and isolates the text area. The configuration panel allows users to choose default settings or customize translations based on language pair and output format preferences. With three output formats available, the app provides flexibility.

For those who prefer typing, an input field is provided, simplifying the translation process. In summary, the Figure Out mobile application stands as an excellent tool for quick and accurate translations on the go. Its user-friendly design and intuitive interface ensure ease of use, while advanced features offer customization and flexibility.

PROJECT EVALUATION AND QUALITY SCENARIO CONTROL

The project evaluation employed the Quantitative Evaluation Framework (QEF), a methodology integrating functionality and user experience assessment to gauge system quality (Stoll et al., 2019; Zheng et al., 2020). Within the QEF framework, the quality scenario is structured into three levels: dimensions, factors, and requirements. The evaluation of the Figure Out System using QEF focuses on three primary dimensions—Functionality (F), Adaptability (A), and Efficiency (E)—aligned with ISO 9126 standards. The Functionality dimension includes four key factors: functional aspects, user interaction, content quality, and connectivity. The Adaptability dimension comprises versatility and maintenance, while Efficiency includes strength, consistency, and integrity (Heidari & Loucopoulos, 2014; Escudeiro & Bidarra, 2008).

Evaluation Results

Throughout the project's duration, there was a noticeable enhancement in the overall average quality of Figure Out, aligning with expectations. This improvement was quantified at each formal evaluation juncture, illustrating the project's progression. Cumulative quality was assessed by dimension, with successive evaluations reflecting an increase in project quality:

- Alpha testing: 76%
- Beta testing: 77%
- Final assessment: 96%

The software testing process spans crucial stages, ensuring the ultimate product's high quality and alignment with user needs. Notably, the incremental quality improvement from Alpha testing to Beta testing, though marginal (76% to 77%), highlights areas for further enhancement in features and functionality. However, the considerable improvements observed from

the Beta version to the final version, particularly in the Functionality dimension (67% improvement), Efficiency (31% improvement), and Adaptability (16% improvement), signify positive strides.

This positive trend suggests a focus on non-functional requirements, prioritizing the creation of a robust platform to support essential functional features. This approach is vital for ensuring user needs are met reliably and efficiently. While significant progress has been made from Beta to the final version, there remains room for improvement. The development team's ongoing focus on both functional and non-functional requirements is crucial to attaining the highest product quality and meeting user expectations. This continuous effort will contribute to the product's success and user satisfaction.

CONCLUSION

Figure Out stands as an innovative assistive technology developed to dismantle communication barriers encountered by everyday users, with a transformative impact on how people communicate and interact. A key advantage lies in its ability to simplify access to information, aiding individuals struggling with foreign languages, especially beneficial for tourists navigating new countries and those who are deaf or hard of hearing.

Notably, Figure Out is tailored to promote inclusion for those relying on sign language as their primary means of communication. Through the integration of national sign languages and International Sign, the technology ensures effective communication for all users, regardless of language or communication preferences. In its continuous development, Figure Out prioritizes enhancing the user experience, including the creation of a visually appealing identity and the integration of additional sign languages to broaden its impact.

In essence, by dismantling communication barriers and fostering inclusion, it contributes to the creation of a more connected and accessible society for all.

ACKNOWLEDGMENT

This work was developed at the R&D GILT with the support of FCT/MCTES (UIDB/05627/2020, UIDB/05627/2020, 950/2019), and Erasmus+ Programme, through project InSign, project number 2019-1-DE01-KA203-004964.

REFERENCES

- Escudeiro, N., Escudeiro, P., Cunha, B., & Gouveia, M. C. (2022). Inclusive cultural heritage tourism. In 4th International Conference on Tourism Technology & Systems (ICOTTS'22), Santiago, Chile. https://doi.org/10.1007/978-981-19-9960-4_19
- Escudeiro, P., & Bidarra, J. (2008). Quantitative evaluation framework (QEF). *Revista Ibérica de Sistemas e Tecnologias de Informação*, 16–27.

- Escudeiro, P., Gouveia, M. C., Cunha, B., & Escudeiro, N. (2022). Accessibility of national cultural heritage to deaf tourists. In Proceedings of the 5th International Conference on Tourism Research, Porto, Portugal. <https://doi.org/10.34190/ictr.15.1.232>
- Escudeiro, P., Escudeiro, N., Reis, R., Barbosa, F. M., Bidarra, J., Baltazar, A., & Gouveia, B. (2013). Virtual Sign Translator. In International Conference on Computer, Networks and Communication Engineering, Beijing, China. <https://doi.org/10.2991/iccnce.2013.72>
- Heidari, F., & Loucopoulos, P. (2014). Quality evaluation framework (QEF): Modeling and evaluating quality of business processes. *International Journal of Accounting Information Systems*, 15(3), 193–223. <https://doi.org/10.1016/j.accinf.2014.01.005>
- House, J. (2016). *Translation as communication across languages and cultures*. Routledge.
- Oliveira, T., Escudeiro, P. M., Escudeiro, N. F., & Rocha, E. (2019a). Automatic sign language translation to improve communication. In *IEEE Global Engineering Education Conference* (pp.937–942), Dubai, UAE. <https://doi.org/10.1109/EDUCON.2019.8725244>
- Oliveira, T., Escudeiro, N., Escudeiro, P., Rocha, E., & Barbosa, F. M. (2019b). The VirtualSign channel for the communication between deaf and hearing users. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 188–195. <https://doi.org/10.1109/RITA.2019.2952270>
- Stoll, S., Camgoz, N. C., Hadfield, S., & Bowden, R. (2019). Text2Sign: Towards sign language production using neural machine translation and generative adversarial networks. *International Journal of Computer Vision*, 128, 891–908. <https://doi.org/10.1007/s11263-019-01157-4>
- Zheng, J., Zhao, Z., Chen, M., Chen, J., Wu, C., Chen, Y., Shi, X., & Tong, Y. (2020). An improved sign language translation model with explainable adaptations for processing long sign sentences. *Computational Intelligence and Neuroscience*. <https://doi.org/10.1155/2020/8816125>