Bridging Accessibility, Innovation, and Multimodal AI for Inclusive Tourism Education

DIMITRIS KOUREMENOS[1,2], KLIMIS NTALIANIS[2], NIKOS E. MASTORAKIS[3,4]
[1]Institute of Informatics and Telecommunications,
NCSR "Demokritos",
Athens,
GREECE

[2]Department of Business Administration,
University of West Attica,
12243 Athens,
GREECE

[3]Technical University of Sofia,
Sofia,
BULGARIA

[4]Hellenic Naval Academy,
Athens,
GREECE

Abstract: - The Greek Language Multimodal Lip Ready (GLaM-Sign) is an innovative prototype multimodal corpus that the Greek research community lacks. It is the first Greek multimodal corpus that combines many modalities, including Greek Sign Language (GSL), in a remarkable volume of 30 hours in media duration with 279,042 words. It is designed for research purposes, can be easily extended, and is primarily intended for training or fine-tuning AI models supporting the inclusivity of society for Deaf and Hard-of-Hearing (DHH) individuals. This corpus was developed as part of the FEELIT project, combining high-resolution speech audio, digital videos of Greek Sign Language, visual lip-reading, and synchronized Greek subtitles. While its primary focus is promoting inclusiveness in Greek tourism, we extend it by adding extra modal lip reading. The primary purpose of this corpus is to be used in the future in other fields like learning hubs, healthcare, and public government services. It is the result of our collective efforts, and now it is our shared responsibility to continue its development. We call upon the AI research community to employ the innovative Greek multimodal corpus prototype in models, contribute to its expansion and improvement, and collaborate with various stakeholders. This would extend it to additional languages and realize a true multimodal and multilingual corpus. The GLaM-Sign shows the significant benefits of inclusive multimodal corpora in Greece, which lacks these vital resources. It would reduce communication gaps, foster innovation, and add value for the next inclusive AI technologies.

Key-Words: - corpus, multimodal, dataset, speech-to-text, sign-language, sign-language-recognition.

Received: March 19, 2024. Revised: December 21, 2024. Accepted: February 14, 2025. Published: May 5, 2025.

1 Introduction

In recent years, many innovative learning hubs have developed to enhance e-learning. The FEELIT project [1] aims to enhance inclusive education within the tourism sector for the Deaf and Hard of Hearing (DHH) tourists. The FEELIT project leverages cutting-edge technologies, such as virtual reality (VR), to certify the skills of the professionals in the tourism industry while addressing the unique needs of Deaf and Hard of Hearing (DHH) tourists, providing them with inclusive tourist experiences.

Tourism supports millions of jobs worldwide (about 10% of the global Gross Domestic Product), [2]. However, DHH tourists face enormous challenges in tourism. The FEELIT project seeks to bridge these gaps by equipping tourism

professionals with the tools and skills to provide accessible and inclusive services, promoting equity and accessibility within the tourism sector.

Apart from FEELIT's project goals, we seized the opportunity to expand the learning materials and create the GLaM-Sign corpus [3], an innovative, groundbreaking, and prototype Greek multimodal corpus designed to advance research and development in Artificial Intelligence (AI), which does not currently exist in Greek reality.

GLaM-Sign corpus includes synchronized multiple modalities of Greek Sign Language digital video, spoken Greek overlay, Greek video lip reading video, and textual synchronized subtitles, supporting the creation and fine-tuning of AI models.

FEELIT focuses on educational training and certification, while GLaM-Sign supports research for DHH inclusion. GLaM-Sign enriches FEELIT's materials with multimodal data, building a long-term foundation for innovation. Together, they effectively support each other.

2 Innovative Features

The FEELIT project's training material is structured in three chapters. The first chapter covers DHH communication and the problems they face, the second chapter introduces the technologies that can enhance DHH accessibility, and the third chapter examines tourists' case studies. FEELIT offers innovative learning practices tailored to the needs of DHH and tourism professionals, promoting the inclusion of DHH tourists.

The FEELIT training material includes video lessons in sign language with text subtitles and over-voiced and visual elements (at least 30 hours or more). The multimodality features of these training materials make them innovative and ensure the learners' knowledge and cognition. The FEELIT training materials are available via a free cloud platform, contributing to the open learning philosophy and making a more inclusive world.

Moreover, aligned with the training material, another innovative learning feature is the development of VR games offering immersive learning experiences. The VR games simulate real-life scenarios of the daily problems the DHH tourists face. The VR games aim to provide tourism professionals with a better understanding of the challenges that DHH tourists face in daily life and enable them to design an accessible tourism product for DHH tourists in the future. Also, we have a desktop version for those who do not have expensive VR equipment. It remains an affordable and inclusive option for learners.

In closing, the FEELIT project introduces interesting, innovative approaches, setting new standards for inclusive learning. Finally, we aim to equip tourism professionals and forthcoming DHH tourist partners with the crucial skills to design a more profitable tourism product for DHH tourists.

3 State of the Art

Although e-learning self-learning systems for DHH students have evolved significantly, several gaps related to accessibility and inclusion remain.

We collected some papers that deal with e-learning for DHH students. Researchers also observe that due to communication difficulties, fatigue is more easily induced after a mental effort, [4]. Learning platforms and educational materials translated into sign language have been developed, demonstrating that the use of sign language significantly increases DHH students' performance in e-learning, [5], [6]. Nevertheless, sign language integration is still a considerable barrier, mainly due to the enormous cost (studio, interpreting, editing), [7]. Similarly, low educational qualifications and sign language are key reasons for the low participation and satisfaction rates of DHH individuals in online courses, [8].

The FEELIT project stands out as it provides high-definition video lessons in sign language, with subtitles and over-voiced and visual material. The GLaM-Sign material is an extension of FEELIT material by extending with an additional video of lip-reading to create a prototype multimodal corpus to be made available to the AI community in order to provide incentives for the development of new or improvement of existing accessibility technologies for DHH individuals. In the next section, we present some interesting research works related to the GSL.

A rule-based machine translation RBMT system was developed to create a GSL glossed corpus and ease the work of specialist translators. A combination of RBMT with statistical machine translation SMT was used to produce a parallel Greek-GSL corpus, getting good results in some areas, such as weather reports, [9]. In the following paragraphs, we will cite the most multimodal corpora works related to lip reading, the Greek language, and its creation purpose. Let us start with a pioneering work in which the MobLip dataset was developed, containing 3,685 Greek words and a total duration of about 30 minutes of video. This dataset aims to train deep-learning models in word

recognition through lip reading, setting a new standard in the field, [10].

Moreover, a practical application of Multimodal Lip-Reading for Tracheostomy Patients was presented, consisting of 4 videos with an estimated duration of 4 minutes each, [11]. The project aims to create an AI model for word recognition for patients with voice loss due to tracheostomy. Moreover, the MUSCLE Movie Database was introduced, including Greek data samples, with a total corpus of 42 minutes of video from English and Greek scenes to recognize dialogues and faces, [12]. Finally, the Bilingual AVASR Corpus (BAVCD) was introduced, containing 2,200 Greek words and video depth recordings of about 35 minutes, aiming to create an application of automatic speech recognition, [13].

The existing multimodal data we mentioned before regarding the Greek language reality presents a lack and limited volume of multimodal data. At the same time, no one incorporates GSL as an additional modality. The FEELIT project, a groundbreaking initiative, come to fill this gap, incorporating Sign Language as an important modality data valuable for the accessibility of DHH. This project offers a remarkable volume of multimodal corpus, more than 30 hours of video, with 279,042 words and 17,989 unique words, and it provides a remarkable corpus with multimodal linguistic data for the Greek scientific AI community.

4 Methodology

The FEELIT project adopted a user-centered design approach, using questionnaires from tourism professionals and DHH people to design and develop accessible training materials and tools. In this way, we ensured that FEELIT's training materials and tools were aligned with DHH's needs.

The questionnaires confirmed, as expected, the main problems DHH people face in tourism and everyday life. The results helped to properly design and develop the FEELIT project tools, which meet the requirements of DHH users. The FEELIT project chose the iterative waterfall model (IWM). IWM combines development in phases that incorporate user feedback, allowing for targeted improvements.

Regarding questionnaires, in the FEELIT project's first phase, data and requirements were collected through questionnaires from 150 people (75 people with DHH disability and 75 tourism professionals). The analysis of this data identified the needs and challenges faced by tourists with DHH. This identification helped to draw valuable conclusions about the requirements for the development of tools and training materials, ensuring that these are tailored to the needs of people with DHH and tourism professionals.

Also, based on the questionnaires, VR simulation scenarios were developed to describe the challenges faced by DHH tourists. These scenarios are intended to help tourism professionals better understand the challenges so they can better serve DHH tourists.

The following sections present detailed findings, where the data will be described in detail, including key demographics, accessibility challenges, and user preferences.

5 Findings and Analysis

5.1 Demographics

In the FEELIT project, we drew useful conclusions that helped us design the systems and training materials. We observed different educational and communication needs among deaf and hard-of-hearing respondents, articulating the need for tailored educational materials for each target subgroup. A total of one hundred fifty participants took part in the study, distributed as equally as possible by disability type and origin. The following Table 1, recaps the FEELIT's participants' profile data:

Table 1. FEELIT's profile

Gender	
Men	50.6%
Women	49.4%
Nationality	
Greeks	37.0%
Italians	24.7%
Romanians	24.7%
Cypriots	13.6%
Education	
Postgraduate degree	13.8%
High school graduates	18.8%
Language	
Sign language	24.7%
Oral communication	24.7%
Lip reading	21.3%

5.2 Travel Preferences

This section will present valuable findings on the preferences and problems DHH tourists face. The findings are intriguing, will be enjoyable for tourist professionals, and are presented in the following Table 2 and Table 3.

Table 2. Respondent Data and Challenges

Category	Statistics/Information
Travel Frequency	1-2 times per year (56.7%)
Travel Companions	with one person (37.8%)
	in groups (32.4%),
	solo (17.6%)
Trip Planning	Online platforms (39.3%)
	friends/family/travel agents
Main Travel Purpose	Leisure/relaxation (26.7%),
	events/cultural experiences
Transportation Methods	Airplanes (26.6%),
	private vehicles (18.4%),
	trains/buses
Less Common Activities	Adventure tourism, cruises, professional trips
Main Barriers	Voice-only airport announcements (38.8%),
	passport controls (16.5%),
	check-ins (14.9%)

Table 2 and Table 3 show that most DHH tourists travel 1-2 times a year, mainly for leisure and to experience a country's culture. They mainly use airplanes as a means of transport, at 26.6%, and private vehicles, at 18.4%. DHH tourists said they have difficulties with verbal announcements at airports, at 38.8%, and passport controls, at 16.5%. Romanian DHH tourists highlighted the need for more accessibility to sign language and visual information. Most DHH tourists use online platforms to book and plan their trips. Once again, we observe the need for inclusion applications using accessible sign language and subtitling.

Table 3. Challenges by Nationality

Nationality	Challenges
Greeks	Voice-only announcements at airports, especially during Deaf-specific events.
Italians	Untrained staff and inadequate accessibility tools abroad
Romanians	Need for sign language interpreters and visual information to reduce confusion.
Cypriots	Fast-speaking officials or unclear announcements.

5.3 Bookings Feedback

We present the questionnaire results on preferences and challenges the DHH tourists face using online resources, transportation, and accommodations for tourist services. The Table 4 illustrates the trends identified.

Table 4. Bookings Feedback

Online Booking Tools and Preferences	
Category	**Statistic/Satisfaction**
Air ticket booking	29.1%
Greek	Prefer internet for ticket booking but rely on traditional travel agencies for some services.
Italian	Appreciate online tools but face issues with untrained airport staff.
Cypriot	Find online booking straightforward, but face communication barriers due to unclear speech from staff.
Ratings of Transportation Hubs and Services	
Airports	Mostly 4- and 5-star reviews
Train and Bus Stations	Moderate evaluations
Ferry Ports	Fewer high ratings
Car Rental Services	Fewer high ratings
Taxi Stands	Mixed levels of satisfaction
Accommodation Type Preferences	
Hotels	54.1%
Guesthouses	14.1%
bed and breakfasts	12.9%)
Rented apartments	50.6%)
camping and caravans	less favored
Staff Behavior	Often rated poorly
Service Availability	Often rated poorly
Facility Quality	Generally rated good or very good.

The survey indicates that online tourism platforms are widely accepted for booking flights and accommodations. DHH tourists are highly satisfied with Airports, while ports and bus/railway stations receive moderate ratings. Hotels were the most preferred accommodation for DHH tourists, although some expressed dissatisfaction with staff behavior and the lack of accessibility.

5.4 Assistive Technologies and Services

Regarding the inclusion technology domain, DHHs participating in the FEELIT survey were asked and presented with the most popular accessible technology solutions. Our findings showed that DHHs prefer integrating visual tools and notifications (visual or written) and using sophisticated accessibility technologies with subtitles or sign language. It also stressed that better training of tourism staff is needed in dealing with DHH tourists. These are the most important points that will significantly improve the tourism experience of DHH tourists.

DHH tourists showed a strong preference for technologies such as subtitling (92%), applications for Deaf tourists (88%), 3D sign language (85%)

(Figure 1), Sign language interpreter services (22%), and video relay services (14%).

DHH participants expressed the very helpful and essential features of accessible technology tools such as real-time captioning and 3D signed translations. Most DHH participants wanted more visual signals on public transport.

Additional technologies such as wearable vibration devices and audio loop systems on public transport (76%) (Figure 1) garnered approval from many respondents who rated them essential or highly useful. We have to provide a broad spectrum of technologies for DHH to help create a more welcoming and enabling environment for travel.

The above preferences of DHH participants show the real needs of DHH tourists and are helpful guides for integrating accessible technologies in tourism services for DHH people.

Finally, we highlight the diverse adapted needs of DHH participants, whether some use lipreading and written text as base communication language and others sign language-based communication. We consider this and develop accessible educational resources and tools adapted to all DHH tourists.

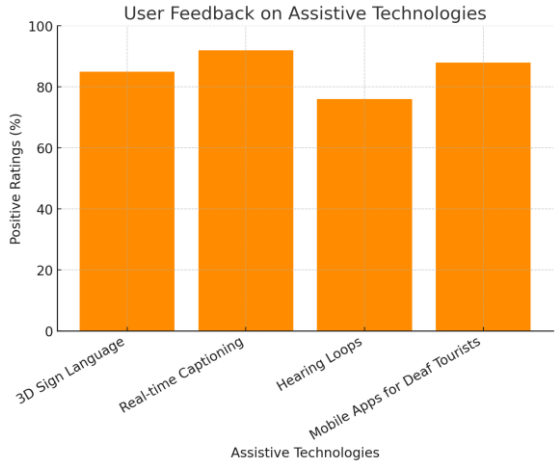

Fig. 1: User Feedback on Assistive Technologies

5.5 Findings Conclusion

The FEELIT DHH participants mentioned their main problems via the questionnaires: communication difficulties, difficulties at airports due to vocal communications, and the lack of sign language interpreters. On the other hand, organized trips with interpreters and nature tourism were evaluated and suggested by DHH visitors as most meaningful and effective. In conclusion, it is proposed to boost support for DHH tourists by adding more SL interpreters, providing trained tourism staff, providing accessible applications (captioned and subtitled information), and improving communication for better accessibility. By implementing the above, policymakers and tourism stakeholders can foster a more inclusive travel experience for DHH tourists.

6 Feelit Educational Material

This section describes the syllabus of FEELIT's training material. The material is structured in three chapters and mentions the accessibility of DHH tourists. It mixes theory and practice and practical technological applications for DHH tourists. The FEELIT aims to train tourism experts to create accessible tourist services for DHH visitors and train DHH adults to be professional tourism cooperators with tourist agents. We will now describe the chapters of FEELIT's training syllabus material.

The FEELIT's texted training material is 35,000 words (120-150 pages) with 50-60 visual elements. Every chapter is split into 4-5 primary sections, and certain sections contain subsections.

In the FEELIT project, we created video lessons using the educational resource with text. The video lessons consist of approximately 50% visual slides on the left side, the video translation into sign language at 30% placed in the right frame, and finally, written subtitles at 20% at the bottom. This arrangement fully meets the needs and preferences of the DHH students (Figure 2), and in the evaluations, we had very high satisfaction. A significant increase of 40% was observed in the DHH learners' understanding and performance of the educational material when it includes visual elements and is translated into sign language, [14]. The same is true in highlighting the positive impact of sign language on the education of DHH learners, [15]. Chapter 1, "Introduction to Tourism," explains the primary categories of tourism. Chapter 2, "Accessible Tourism," examines universally accessible design in tourism services. Finally, Chapter 3, "Technology and Support," analyzes accessible technological tools, digital sign language, and other accessible technologies.

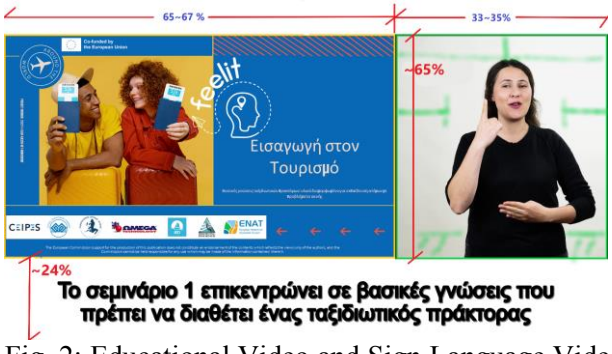

Fig. 2: Educational Video and Sign Language Video Layout

The FEELIT training material is fully adapted, accessible, and translated into sign language, with subtitles and a voice-over, and offers a legend to tourism professionals to improve the DHH visitor experience.

7 Greek Multimodal Corpus and AI Applications for Accessibility

In this section, we will analyze the potential applications and uses of the promising multimodal GLaM-Sign corpus. This research is based on the educational materials from the FEELIT project, although it was not part of the project's obligations and duties. The main goal of GLaM-Sign is to address a significant gap in the Greek context. This multimodal material is remarkable because it includes the typical modalities such as video, speech, and subtitles, and it incorporates Greek Sign Language in the video (Figure 2, Figure 3, Figure 4).

The GLaM-Sign corpus is stored in the cloud repository Hugging Face platform and is openly available. They will be used to pursue future synergy with interested researchers or students to develop accessible AI tools adapted to DHH inhabitants. Using the GLaM-Sign corpus, we can develop numerous innovative AI applications for DHH persons. Such as automatic speech recognition (ASR) of Greek by fine-tuning existing models, automatic speech recognition through lip reading in noisy environments, sign recognition, and sign language generation (sign recognition and text-to-sign), as well as advanced solutions such as chatbots or hotel bots that communicate directly in sign language. In addition, immersive applications, such as virtual assistants and virtual guides for museums, are just a few examples that demonstrate the almost unlimited potential uses for exploiting this corpus.

It is known that Greek-language multimodal datasets remain underrepresented, especially those integrating Greek Sign Language (GSL) with synchronized modalities like speech, lip video, and subtitles.

GLaM-Sign fills this gap by offering a native Greek multimodal resource that could promote developing Greek ASR models in noisy environments, incorporating lip reading and audio, Text-to-Sign Language generation models, and End-to-end multimodal transformers for dialogue systems in sign language. Moreover, we could fine-tune existing open models or frameworks, such as OpenAI's Whisper (audio-to-text) [16], LipNet (lip reading) [17], and Sign2Text, using a Deep Learning-based Sign Language Translation System [18].

The absence of a Greek corpus incorporating all these modalities has hindered the development of localized accessibility tools. GLaM-Sign offers a new benchmark for developing and evaluating such models within the Greek context.

Fig. 3: Video Lip Reading

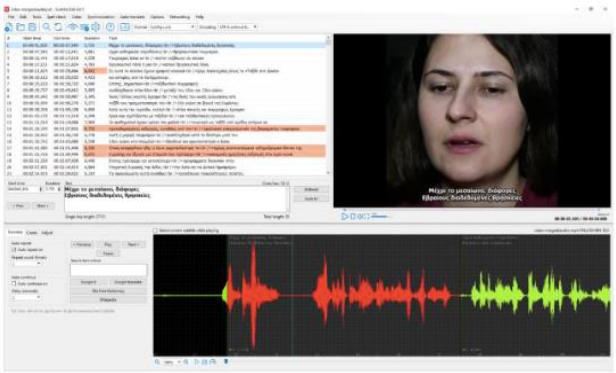

Fig. 4: Speech subtitles word-level timestamps synchronization

Finally, we hope that the GLaM-Sign corpus will find a warm response from the research community, not only for developing innovative, accessible applications needed by the deaf community but also for expanding the corpus into other thematic areas. Our goal is the shared vision of all of us to achieve a more accessible world without language barriers for DHH people.

8 Conclusion

The FEELIT project and the GLaM-Sign corpus mark an important milestone. By integrating multimodal resources (speech, GSL, lip reading cues, subtitles), we offer a valuable learning hub and a research corpus that opens up new possibilities for

artificial intelligence models and people with hearing impairments (DHH).

The multimodal materials from the FEELIT project focus on and cover the whole of tourism. Therefore, tourism-related accessibility applications may be developed, such as interactive translation applications that provide real-time sign language interpretation at hotel check-ins or virtual assistants using sign language avatars. This multimodal corpus, which is ready to redefine the travel experience of DHH visitors, can have many applications. We invite the academic and AI communities to build upon this corpus by refining existing models or developing new architectures for sign language understanding, generation, or multimodal interaction.

In conclusion, the FEELIT project and the multimodal corpus prove that even educational Erasmus projects that aim to produce training material can form the basis for developing the multimodal synchronized and annotated corpus that the AI community needs. Our vision is to empower AI experts, the academic community, and the Deaf community to create a more accessible environment using real-world AI tools.

References:

[1] FEELIT Project, "FEELIT: Advancing Accessibility and Inclusion in Tourism", [Online]. https://feelit.infoproject.eu. (Accessed Date: January 30, 2025).

[2] World Travel & Tourism Council, "Economic Impact Reports 2023: Global Trends in Travel and Tourism," *World Travel & Tourism Council*, 2023, [Online]. https://wttc.org/Research/Economic-Impact (Accessed Date: January 30, 2025).

[3] D. Kouremenos and K. Ntalianis, "GLaM-Sign: Greek Language Multimodal Lip Reading with Integrated Sign Language Accessibility," *ArXiv Prepr. ArXiv250105213*, Jan. 2025, https://doi.org/10.48550/arXiv.2501.05213

[4] N. A. Hashim, N. Safie, and M. Mukhtar, "A review of e-learning models for deaf and hearing impaired people," *J Eng Appl Sci*, vol. 13, no. 21, pp. 9029–9037, Nov. 2018, doi: 10.3923/jeasci.2018.9029.9037.

[5] N. Krishnamoorthy, A. Raveendran, P. Vadiveswaran, S. R. Arulraj, K. Manathunga, and S. Siriwardana, "E-Learning Platform for Hearing Impaired Students," in *2021 3rd International Conference on Advancements in Computing (ICAC)*, Dec. 2021, pp. 122–127. doi: 10.1109/ICAC54203.2021.9671113.

[6] M. H. Mohd Hashim and Z. Tasir, "An e-learning environment embedded with sign language videos: research into its usability and the academic performance and learning patterns of deaf students," *Educ. Technol. Res. Dev.*, vol. 68, no. 6, pp. 2873–2911, Dec. 2020, doi: 10.1007/s11423-020-09802-4.

[7] P. Martins, H. Rodrigues, T. Rocha, M. Francisco, and L. Morgado, "Accessible Options for Deaf People in e-Learning Platforms: Technology Solutions for Sign Language Translation," *Procedia Comput. Sci.*, vol. 67, pp. 263–272, Jan. 2015, doi: 10.1016/j.procs.2015.09.270.

[8] E. F. Lago and S. O. Acedo, "Factors affecting the participation of the deaf and hard of hearing in e-learning and their satisfaction: A Quantitative study," *Int. Rev. Res. Open Distrib. Learn.*, vol. 18, no. 7, Nov. 2017. https://doi.org/10.19173/irrodl.v18i7.2862.

[9] D. Kouremenos, K. Ntalianis, and S. Kollias, "A Novel Rule Based Machine Translation Scheme from Greek to Greek Sign Language: Production of Different Types of Large Corpora and Language Models Evaluation," *Comput. Speech Lang.*, vol. 51, pp. 110–135, 2018. https://doi.org/10.1016/j.csl.2018.04.001.

[10] T. Exarchos, G. N. Dimitrakopoulos, A. G. Vrahatis, G. Chrysovitsiotis, Z. Zachou, and E. Kyrodimos, "Lip-Reading Advancements: A 3D Convolutional Neural Network/Long Short-Term Memory Fusion for Precise Word Recognition," *BioMedInformatics*, vol. 4, no. 1, Art. no. 1, Mar. 2024, doi: 10.3390/biomedinformatics4010023.

[11] Y. Voutos, G. Drakopoulos, G. Chrysovitsiotis, Z. Zachou, D. Kikidis, E. Kyrodimos, and T. Exarchos, "Multimodal Lip-Reading for Tracheostomy Patients in the Greek Language," *Computers*, vol. 11, no. 3, Art. no. 3, Mar. 2022, doi: 10.3390/computers11030034.

[12] D. Spachos, A. Zlantintsi, V. Moschou, P. Antonopoulos, E. Benetos, M. Kotti, K. Tzimouli, C. Kotropoulos, N. Nikolaidis, and P. Maragos, "MUSCLE movie-database: A multimodal corpus with rich annotation for dialogue and Saliency Detection," in *6th Language Resources and Evaluation Conference*, Marrakech, Morocco, May 2008, pp. 16–19, [Online].

https://openaccess.city.ac.uk/id/eprint/2771/ (Accessed Date: January 30, 2025).

[13] G. Galatas, G. Potamianos, D. Kosmopoulos, C. McMurrough, and F. Makedon, "Bilingual corpus for AVASR using multiple sensors and depth information," in *Auditory-Visual Speech Processing*, Volterra, Italy, Sep. 2011, pp. 103–106, [Online]. https://www.isca-archive.org/avsp_2011/galatas11_avsp.html (Accessed Date: February 15, 2025).

[14] E. A. C. Carpenter, A. Meltzer, and M. Marquart, "Best Practices for Inclusivity of Deaf/deaf/Hard of Hearing Students in the Synchronous Online Classroom," *World J. Educ.*, vol. 10, no. 4, Art. no. 4, Aug. 2020, doi: 10.5430/wje.v10n4p26.

[15] M. E. Skyer, "Multimodal Transduction and Translanguaging in Deaf Pedagogy," *Languages*, vol. 8, no. 2, Art. no. 2, Jun. 2023, doi: 10.3390/languages8020127.

[16] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Whisper: Robust Speech Recognition via Large-Scale Weak Supervision," *ArXiv E-Prints ArXiv221204356*, Dec. 2022, https://doi.org/10.48550/arXiv.2212.04356.

[17] Y. M. Assael, B. Shillingford, S. Whiteson, and N. De Freitas, "Lipnet: End-to-end sentence-level lipreading," *ArXiv Prepr. ArXiv161101599*, 2016, https://doi.org/10.48550/arXiv.1611.01599.

[18] G. L. P and R. Francis, "Sign2Text: Deep Learning-based Sign Language Translation System Using Vision Transformers and PHI-1.5B," in *2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, Kota Kinabalu, Malaysia, Aug. 2024, pp. 282–287. doi: 10.1109/IICAIET62352.2024.10730290.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

All the authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution and reviewing the article. No ghostwriter was involved, and no contributor has been left out.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

This research was supported by funding from the Erasmus+ program under the project 'FeelIt: 2021-1-EL01-KA220-VET-000028082'. For more information, visit https://feelit.infoproject.eu.

Conflict of Interest

The authors have no conflicts of interest to declare.