



HAL
open science

Special issue on sign language translation and avatar technology

Rosalee Wolfe, Annelies Braffort, Eleni Efthimiou, Evita Fotinea, Thomas Hanke, Dimitar Shterionov

► To cite this version:

Rosalee Wolfe, Annelies Braffort, Eleni Efthimiou, Evita Fotinea, Thomas Hanke, et al.. Special issue on sign language translation and avatar technology. Universal Access in the Information Society, 2023. hal-04221882

HAL Id: hal-04221882

<https://hal.science/hal-04221882>

Submitted on 24 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Special issue on sign language translation and avatar technology

Rosalee Wolfe¹, Annelies Braffort², Eleni Efthimiou¹,
Evita Fotinea¹, Thomas Hanke³ & Dimitar Shterionov⁴

1. Institute for Language and Speech Processing, Athens, Greece
2. Université Paris-Saclay, Orsay, France
3. Universität Hamburg, Hamburg, Germany
4. Tilburg University, Tilburg, The Netherlands

Deaf communities around the world face barriers in accessing education, health care, job opportunities and social services. Many hearing people unfamiliar with the daily challenges faced by these groups are surprised to learn that the challenges arise not primarily from a barrier of sound, but from a barrier of language. This is particularly true for those who are born deaf or become deaf at an early age.

The first or preferred language of early deafened people is a signed language, not a spoken one. Instead of utilizing the oral/aural modality of spoken language, signed languages use a gestural/visual modality for producing and receiving language. Signed languages, such as American Sign Language (ASL), French Sign Language (LSF), German Sign Language (DGS) and Greek Sign Language (GSL), are unique, autonomous, fully complete natural languages that are distinct from any ambient spoken language as well as from each other. In other words, ASL is different from English; LSF is different from French, etc., but also ASL is different from GSL, LSF, DGS and other signed languages. In fact, linguists have identified over two hundred signed languages, and the list continues to grow [1]. With their own unique grammars, signed languages are fully capable of the same range of subtlety, precision, power and ability to express complex thoughts as a spoken language [2].

For communities that use a signed language as their first language, any spoken language *in any form* is a second language. Literacy of written language for these groups is very low. For example, in the USA, the reading level of deaf students graduating high school is between the third and fourth grade [3]. For this reason, Deaf¹ sign language users around the world face continual challenges in daily interactions with hearing, non-signing populations.

The language barrier causes difficulties in accessing health care, education and job opportunities as well as legal consultation. For these critical services, the gold standard for facilitating communication has always been and will always be engaging the services of a certified sign language interpreter. However, in daily life, there are many short conversations, such as those at a store counter, over a hotel desk or in an office foyer. These interactions are so short in duration that hiring a qualified interpreter would be prohibitively expensive, or even impossible because in most countries there is a shortage of qualified interpreters. For quotidian interactions, an automatic translation system between spoken and signed language could remove communication obstacles and improve inclusivity. For technology of this sort to be effective, it must produce sign language in a way that is acceptable to members of the sign language community.

1 The state of the art

There has been significant progress in machine translation of spoken languages, to the point where it is widely accepted for use in many small everyday situations [4]. So, at first glance, it might seem that the goal of improving communication between Deaf and hearing communities is at our metaphorical doorstep. However, when compared to automatic translation between spoken languages, the state of the art for automatic translation between signed and spoken languages lags far behind.

Two major challenges contribute to this slower development, the first of which is the difficulty of representing signed language in a way that is amenable to processing through automatic means. Machine translation techniques for spoken languages employ various strategies to encode (recognize and understand) source language and decode (synthesize or generate) corresponding utterances in a target language, but one thing they have in common is their reliance on text to represent language. Unlike spoken language, which can be expressed as a linear stream of text, signed languages entail linguistic processes that cannot be linearly ordered because their linguistic system exploits the properties of visual–gestural channels: A lot of information is expressed simultaneously and organized in space, and iconicity plays a central role [5]. How best to represent sign language for machine translation is still an open question. Some approaches to representing sign language, such as motion capture, contain so much detail that it is prohibitive to use them to generate new signed utterances [6]. At the other extreme, there are representations, such as glosses that are so parsimonious that they are missing data essential for effective language sign generation.

The second major challenge is a lack of an avatar technology that effectively displays generated sign language in a manner that is legible and acceptable to end users. This might seem quite puzzling since there has been over a century’s worth of development in animation for movies and fifty years of avatar development for video game applications. A casual observer might be tempted to think that using avatars to display sign language is a solved, or nearly solved problem; however, users have been consistent in expressing their concerns about the quality since the appearance of the first signing avatars in the late 1990s [7], 8, 9]. For effective display of sign language, avatars not only must have the communicative power of animated movies that require thousands of hours of manual labor, but also offer the flexibility of real-time responsiveness of a video game and be highly interactive. Until researchers successfully address this challenge, there is little chance that machine translation between signed and spoken languages will become an accepted technology among Deaf communities.

Portraying sign language as animation does have similarities with animation created for movies, but the requirements are higher. Portraying sign language requires more precision in position and moving body joints, especially the hands. This motivated researchers to develop more accurate modeling of the behavior of hand bones, particularly the base joint of the thumb, and building intelligent collision avoidance strategies. Avatar efforts in the early 2000s concentrated on automated portrayal of selected linguistic processes, not yet as a coherent whole, but as demonstration samples. This included examples of fingerspelling and conjugating agreement verbs [10].

2 Establishing SLTAT

Toward the goal of better Deaf–hearing communication through noninvasive technology, several researchers came together in 2011 to establish a workshop called “Sign Language Translation and Avatar Technology” (SLTAT). The workshop convenes on an annual or semiannual basis to discuss recent innovations with like-minded researchers and share recent results. Attendees have backgrounds in multiple disciplines with skills in such areas as linguistics, machine translation, animation, kinematics, native knowledge of multiple sign languages, user experience and Deaf culture.

The exchange of ideas across disciplines has fostered new advances in supporting the automatic portrayal of signed language. Examples include the development of portraying sign language classifiers, which are highly visual and have no easy translation into words of spoken languages [11], and the portrayal of multiple processes that occur simultaneously on a signer’s face [12]. Increased efforts in involving Deaf researchers and Deaf users have resulted in better authoring tools [13] and more Deaf-friendly methods for evaluation [14]. The results have facilitated a steady improvement in the legibility of the sign languages portrayed via avatar [15].

3 In this issue

It was the seventh edition of SLTAT, held in Marseille on June 25, 2022, which inspired this special issue. This issue contains papers that address the question of sign language translation, avatar display and their application.

Two of the papers focus specifically on improving the quality of avatar signing. In *Multi-target Modeling of Spine Movements for Signing Avatar*, Shatabdi Choudhury describes an analysis of signers' spinal motion to build a predictive model of spinal movement based on the signers' wrist positions. This approach has the potential for easing the labor-intensive process of manual animation and automatically reducing the robotic nature of typical avatar motion. In *Considerations on Generating Facial Nonmanual Signals on Signing Avatars*, John McDonald describes an innovative approach to modeling facial animation that facilitates an improved portrayal of the timing and intensity of independent co-occurring linguistic processes that appear simultaneously on a signer's face.

Two additional papers describe complete end-to-end systems. In the first, *Neural Machine Translation from Text to Sign Language*, José De Martino et al. describe a translation system that accepts Brazilian Portuguese text and produces animations of Brazilian Sign Language (Libras). They are currently focusing on translating reading materials for Deaf elementary school children, but plan to increase their scope of work. The paper *Greek Sign Language Recognition for an Education Platform* by Katerina Papadimitriou et al. also describes an educational application, but this application is for hearing students learning sign language. The focus of this paper is a portion of their end-to-end system that supports automatic assessment of student progress in learners' productions of sign language through the application of sign language recognition.

The paper, *Best Practices for Sign Language Technology Research* by Fox et al., gives advice on effective strategies for researchers developing new sign language technologies. The authors emphasize the centrality of Deaf communities to any effort of this kind. They urge (hearing) researchers to learn a sign language and become involved with the local Deaf community, and to involve Deaf members in all aspects of the research. Adhering to user-centered design principles such as these maximizes the chances of an outcome that will prove useful in improving communication between Deaf and hearing communities.

4 Conclusion

This special issue on sign language translation and avatar technology documents recent advances toward better Deaf/hearing communication through noninvasive means. We hope you find this special issue informative and useful.

Notes

1. The term "deaf" (little-d deaf) refers to the condition of not hearing. The term "Deaf" (big-D Deaf) refers to people who identify with a community that shares a signed language as their preferred, or first language and has its own culture and sense of identity based on that shared language.

References

1. "Welcome to the SIGN-HUB platform," 2020. [Online]. Available: <https://thesignhub.eu/>. (2023) Accessed 25 May 2023
2. Valli, C., Lucas, C.: *Linguistics of American sign language: an introduction*. Gallaudet University Press, Washington D.C. (2000)
3. Traxler, C.B.: The Stanford achievement test: national norming and performance standards for deaf and hard-of-hearing students. *J. Deaf Stud. Deaf Educ.* 5, 337–348 (2000)

4. Webster, R., Fonteyne, M., Texcan, A., Macken, L., Daems, J.: Gutenberg goes neural: comparing features of dutch human translations with raw neural machine translation outputs in a corpus of english literary classics. *Informatics* 7(3), 32 (2020)
5. Filhol, M. and Braffort, A.: What constraints for representing multilinearity in sign language. In: *Constraint Solving and Language Processing*, Orléans, France (2012)
6. Wolfe, R., McDonald, J.C., Hanke, T., Ebling, S., Van Landuyt, D., Picron, F., Krausneker, V., Efthimiou, E., Fotinea, E., Braffort, A.: Sign language avatars: a question of representation. *Information* 13, 206–222 (2022)
7. Verlinden, M., Tijsseling, C. and Frowein, H.: A signing avatar on the WWW. In: *International Gesture Workshop*, (2001)
8. Krausneker, V. and Schügerl, S.: Avatars for sign languages: best practice from the perspective of deaf users. In: Petz, A., Hoogerwerf, E. J., and Mavrou, K. (eds.) *Assistive Technology, Accessibility and Inclusion, ICCHP-AAATE 2022 Access Compendium*, Linz, Johannes Kepler University, (2022)
9. Angelini, R.: Contrasting technologists' and activists' positions on signing avatars. In: *Extended Abstracts of CHI Conference on Human Factors in Computing Systems*, Hamburg, (2023)
10. Toro, J.: Automated 3D animation system to inflect agreement verbs. In: *Sixth High Desert Linguistics Conference*, Albuquerque, NM, (2004)
11. Filhol, M. and McDonald, J.: The synthesis of complex shape deployments in sign language. In: *Proceedings of the 9th Workshop on the Representation and Processing of Sign Languages*, Marseille, France (2020)
12. Schnepf, J., Wolfe, R., McDonald, J. and Toro, J.: Combining emotion and facial nonmanual signals in synthesized American sign language. In: *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '12)*, Dundee, UK, (2012)
13. Filhol, M.: *Modélisation, traitement automatique et outillage logiciel des langues des signes [Doctoral dissertation]*. Université Paris-Saclay, Gif-sur-Yvette, France (2021)
14. Efthimiou, E., Fotinea, S.-E., Goulas, T., Vacalopoulou, A., Vasilaki, K., Dimou, A.-L.: Sign language technologies and the critical role of SL resources in view of future internet accessibility services. *Technologies* 7(1), 18 (2019)
15. Dimou, A.-L., Papavassiliou, V., McDonald, J., Goulas, T., Vasilaki, K., Vacalopoulou, A., Fotinea, S. E., Efthimiou, E. and Wolfe, R.: Signing avatar performance evaluation within easier project. In: *Seventh International Workshop on Sign Language Translation and Avatar Technology: The Junction of the Visual and the Textual: Challenges and Perspectives*, Marseilles, France, (2022)