



OPEN ACCESS

EDITED BY

Adam Schembri,
University of Birmingham,
United Kingdom

REVIEWED BY

Angus B. Grieve-Smith,
The New School, United States
Maartje De Meulder,
HU University of Applied Sciences
Utrecht, Netherlands

*CORRESPONDENCE

Athanasia-Lida Dimou
ndimou@athenarc.gr

SPECIALTY SECTION

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

RECEIVED 20 October 2021

ACCEPTED 11 July 2022

PUBLISHED 01 August 2022

CITATION

Dimou A-L, Papavassiliou V, Goulas T,
Vasilaki K, Vacalopoulou A, Fotinea S-E
and Efthimiou E (2022) What about
synthetic signing? A methodology for
signer involvement in the development
of avatar technology with generative
capacity. *Front. Commun.* 7:798644.
doi: 10.3389/fcomm.2022.798644

COPYRIGHT

© 2022 Dimou, Papavassiliou, Goulas,
Vasilaki, Vacalopoulou, Fotinea and
Efthimiou. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

What about synthetic signing? A methodology for signer involvement in the development of avatar technology with generative capacity

Athanasia-Lida Dimou*, Vassilis Papavassiliou,
Theodoros Goulas, Kyriaki Vasilaki, Anna Vacalopoulou,
Stavroula-Evita Fotinea and Eleni Efthimiou

Embodied Interaction and Robotics, Institute for Language and Speech Processing (ILSP)/Athena Research Center (ATHENA R.C.), Maroussi, Greece

Although signing avatar technology seems to be the only option currently available to serve sign language (SL) display in the context of applications which demand generative capacity from the part of the technology like in machine translation to SL, signing avatars have not yet been accepted by signers' communities. One major factor for this rejection is the feeling that technology is developed without the involvement of its actual users. Aiming to invite the signers' community into the process of signing avatar development, we have designed the shell methodological framework for signer-informed technology which is implemented as on-line surveys addressed to signer communities of different SLs. The surveys are communicated *via* focused on-line questionnaires with content of signing avatar performance that allows rating of various aspects of the produced SL synthetic signing by human signers. Here we report on the first survey application with content from the Greek Sign Language (GSL). The analysis of the obtained results is 2-fold: it highlights the significance of signer involvement and the provided feedback in the technological development of synthetic signing; in parallel it reveals those aspects of the survey setup that need fine-tuning before its next distribution cycles. The implementation of the first on-line survey can be found in: <https://sign.ilsp.gr/slt-survey/>.

KEYWORDS

signing avatar performance, on-line survey framework, signing avatar rating, avatar generative capacity, user acceptance, signer involvement

Introduction

Traditionally, sign languages have been languages which are developed and transferred within the deaf¹ communities of the world. Since SLs are articulated in the three-dimensional space on and in front of the signer's body, video has been

1 Although the standard academic reference d/Deaf is extensively used in the literature as in [Kusters et al. \(2017\)](#), "deaf" is used here as a generalized term to refer to deaf signers, adopting a recent recommendation by the European Union of the Deaf (EUD) (<https://www.eud.eu/>).

extensively used as a means of SL representation ever since this technology has been made available. The significance of video technology for the representation of signed content also becomes obvious if one considers the lack of a universally accepted system for the written representation of SLs. The use of video has also opened the way to creating sign language corpora, thus enabling the corpus-based linguistic analysis of sign languages and the study of the three-dimensional articulation of signing *via* simultaneous engagement of various articulators (hands, upper body, head, face) on a completely new basis.

However, beyond the revolution that video has brought with respect to SL representation, its use as a means of communication *per se* has several limitations. These include difficulty in editing and lack of signer anonymity. Furthermore, good video quality requires studio infrastructure with proper lighting, capturing devices and video processing equipment, which is neither affordable nor easy to set up. These parameters make it impossible to use video in a manner that would resemble the production of a written document by means of an editing tool or allow easy production of signed messages which, for example, display the result of an automatic translation application like google translate, or modification of a saved file of signing content, as is the case with standard word processors for spoken languages.

In the last two decades, various research attempts to address these problems have focused on developing dynamic sign language representation engines, which use avatar characters for the display of signed utterances. These aims to provide tools that can be used to easily compose, save, modify, and reuse SL content in various education and communication settings, where this lack creates barriers for deaf and hard-of-hearing individuals. Since the earliest attempts for synthetic signing representation, however, deaf communities have received avatar signing with skepticism, if not with complete rejection (Sáfár and Glauert, 2012; Erard, 2017; European Union of the Deaf, 2018; World Federation of the Deaf, 2018). This is partially due to the immature technology underlying SL display, and partially because the signer communities have been kept away from the technology development procedures underlying the synthetic signing performance, which was indeed far from acceptable to technology consumers.

While avatar characters have been developed to be used mainly in gaming, their use was gradually extended to other applications, from film making to various education and communication tools. Sign language synthesis² researchers have identified the opportunity to use avatars for the display of signing since the '90s. However, since the very beginning they

have recognized the demand to develop enhanced display engines that could provide fine-grained motion capabilities as regards the hands and fingers, the body, and the head, as well as advanced expressivity with respect to all face features, and that SL display demand extends far beyond the capabilities of commonly used avatar characters. This demand has been driving dynamic signing avatar technology ever since, and it still poses numerous questions affecting research in the fields of both technology development and SL theoretical analysis, targeting a systematic approach to the incorporation of SL articulation features in synthetic signing environments. Thus, the combination of technological enhancements and SL analysis has proven necessary in view of reaching synthetically composed SL messages that can be recognized as both comprehensible and close to human in respect to signing performance.

In the next sections, we attempt to provide an answer to the question of signing avatar acceptability by showcasing the importance of end-user involvement in the development of the technology. To do so, we developed a methodological framework for involving signers' community in technology development which makes use of a shell environment offered in the form of an on-line questionnaire with the aim to reach as many end-users as possible. The questionnaire underlying structure allows for accommodation of content from different SLs and may emphasize on various aspects of SL articulation. This allows for an iterative process of signers' consultation accommodating content which each time displays the state of technology to be evaluated, thus providing a way of steady communication with and involvement of a wide group of signers during technology development. The goal here is that active involvement of users by means of steady inspection and evaluation of the produced synthetic signing can set the guidelines for the next research goals, while effectively making users participate in the formulation of SL display engines.

Next, we report on the first application of the on-line questionnaire to get user feedback in a survey that intended to identify end-user preferences as regards a specific set of sign articulation features already implemented in two well-known SL avatar technologies. Each of these two technologies represents over two decades of development based on advancements in theoretical research in the last 60 years. These engines, with constant user involvement, may get closer to what signers regard as legibility of the synthetic signed message. In Section Technological and societal background, we present a brief review of the current state of synthetic signing technologies, as well as some societal factors that have incited reticence on behalf of the deaf communities toward avatar signing so far. In Section Methodological framework of the on-line survey application, we present our approach to directly involving deaf individuals in the development of avatar display technology for SL. This is accomplished *via* a series of surveys which use specially designed on-line questionnaires to collect signer preferences regarding various aspects of avatar performance. In Section Results,

² Although several methodologies have been developed to display SL from sources like video or motion capture data using avatars, here we refer exclusively to signing avatar technologies that can support the dynamic composition of new signed messages in contexts such as SL machine translation or SL editors.

we discuss the results of the first application of the survey, addressing the Greek Sign Language (GSL) community. Finally, in Section Discussion, we provide an overview of the experience gained and our future steps toward further addressing other European SL communities in a steady attempt to enhance the generative capacity of the EASIER SL display engine.

Technological and societal background

Currently, SL display technology can be classified along two distinct dimensions: (a) the appearance of the virtual signer, and (b) the motion of the virtual signer (Wolfe et al., 2021). The quality of these two aspects forms the baseline that determines the degree of user acceptance of the SL message produced by the technology.

A virtual signer's appearance can be of three forms: (i) a video recording of a human signer, (ii) a cartoon character, or (iii) a 3D avatar. As mentioned, although video recordings provide the highest degree of realism, they are extremely difficult to edit (Schödl et al., 2000), despite recent advances toward video reuse (Radford et al., 2016). A cartoon character is a simplified representation of a human form, with the details stripped away to increase its communicative power. But although cartoon characters have been used successfully with children (Adamo-Villani et al., 2013), adults prefer more realism to serve their communicative needs³ (Kipp et al., 2011). In addition, a 3D avatar offers more realism than a cartoon character and has the flexibility to display signed messages without the need for pre-recorded video (Jennings et al., 2010; Wolfe et al., 2011; Pauser, 2019). In general, appearance should be determined by the targeted audience and communicative setting.

The motion of a virtual signer is far more significant than its appearance (Krausneker and Schügerl, 2021; Wolfe et al., 2021). In SL, the correct motion issue is critical, because motion is essential to understanding the language; when the motion is wrong, the message becomes difficult or impossible to comprehend.

There are three alternative technologies to drive a signing avatar's motion: 1. Motion capture (mocap); 2. Keyframe animation; 3. Procedural animation. Let us briefly go through these technologies. 1. The data acquired by mocap is human motion, which is recorded *via* markers that are placed on the signer's body and face (Brun et al., 2016), or recognized from video by computer vision techniques. Although the motion data obtained in this fashion is natural in quality, this alternative, which is similar to video, lacks generative capacity. The movement in a pre-recorded phrase is natural, but to create a new phrase from it often tends to be very difficult since an

enormous amount of resources is required, while the result is not necessarily satisfactory. 2. Keyframe animation is based on the observation of motion in natural signing to communicate its salient features through animation software (Wolfe et al., 2011). The result is a library of signs, or sign fragments, which can be easily used to create new phrases. 3. Procedural animation is an avatar-based technology which creates synthetic signing automatically based on linguistic representations of SL (Jennings et al., 2010) corresponding to a library of motions. The last two techniques have been inspired, in their early steps, by work in speech synthesis. The main working hypothesis is that, if we decompose the signs of a SL in articulation segments and create a library of motions that feeds an avatar with these segments, we can generate the synthetic representation of any sign which is composed of pieces of articulation in the library. Each of the avatar animation technologies mentioned has limitations. Keyframe and procedural avatar animation require considerably fewer computing resources to generate new phrases. But although procedural avatar animation technology is potentially the most powerful one in creating new phrases, the motion in the phrases can look "stiff," "awkward," or "hard to read," according to user assessments.

Recent research attempts to generate new phrases from previously recorded video *via* generative adversarial networks (GANs) (Stoll et al., 2018) have yet a long way to go toward producing results that can successfully portray all aspects of SL including accurate handshapes as well as linguistic and affective processes that co-occur on a signer's face. This is not entirely surprising as machine learning approaches require an extensive number of examples from which a neural network can learn, and unfortunately, the amount available SL data is miniscule compared with the amount available for spoken language translation. Further, due to the over 600 skeletal muscles attached in various ways to 206 bones in the human body, the amount of data necessary to demonstrate every motion in SL would be prohibitive. Currently, this technology is not mature enough to create new SL phrases that are comprehensible. This leaves us with keyframe and procedural avatar animation systems for dynamic synthetic signing.

Previous efforts have placed emphasis on the avatar's appearance, but less attention has been paid to the way the avatar moves (Krausneker and Schügerl, 2021; Wolfe et al., 2021). The quality of the motion is essential for comprehension as SL involves a lot of processes interacting in concert on various parts of the body. The human body moves in coordinated, but asynchronous ways, for example, the eyes and head move before the torso, and the torso will tend to move before the arms (McDonald et al., 2016). Eyebrows can express happiness (up movement) and a WH-question (down movement) simultaneously.

Previous avatars were limited to one scripted motion on each body part and so had to be scheduled sequentially. It was impossible to accurately schedule co-occurring linguistic events

³ A thorough review of the differences between cartoon animation and avatar technology is provided in Wolfe et al. (2021).

(Wolfe et al., 2021) due to limitations in their representations. Furthermore, aspects of SL articulation that are often omitted from avatars include mouth gestures, mouthing, and affect. Thus, the motion representations used as input to the signing avatar are of critical importance. This makes quality of motion regarding the whole set of SL articulators one of the major characteristics that distinguish signing avatars from video game avatars and other computer-generated humanlike characters, also being decisive for the comprehensibility of the avatar's signing.

On the other hand, very little has been reported so far with respect to evaluation of avatar comprehensibility by end-users (Kipp et al., 2011). The best published results to date put comprehension rates at 52% (Pauser, 2019), which, however, is not sufficient for effective communication.

On the user side, lack of successful motion articulation, or even worse, lack of SL representations which incorporate features of the simultaneous multilayer articulation of natural signing, has been one of the main reasons which have led deaf communities to reject SL avatars (Sáfár and Glauert, 2012; Erard, 2017; European Union of the Deaf, 2018; World Federation of the Deaf, 2018). To date signing avatars presented as working solutions to signers' community "have exhibited robotic movement and are mostly unable to reproduce all of the multimodal articulation mechanisms necessary to be legible, comparable to early speech synthesis systems which featured robotic-sounding voices that chained words together with little regard to coarticulation and no attention to prosody" (Wolfe et al., 2021). Display of sign language requires precision in communicative power, to be able to achieve the required comprehensibility and naturalness in signing, which would make it acceptable by human signers.

Apart from robotic motion, there are additional factors that have fueled the negative attitude of deaf communities toward signing avatars. In many cases, deaf signers identify themselves as members of a minority group to which language is the main carrier of cultural heritage and identity, rather than persons with a disability (De Meulder et al., 2019). As such, they need to continuously struggle with policy making issues on local, national, or even global level to establish their right to use their SL for all communication purposes and have the right to face a hearing majority (Branson and Miller, 1998) who is not familiar with SL user communities' reality. Thus, barriers of distrust are also added to the language barrier between deaf and hearing communities. In this context, various already proposed machine translation systems that exploit signing avatars for the display of the signed translation output are far from satisfactory. Thus, instead of been viewed as an assistive tool, the technology is perceived as an unsatisfactory replacement for human interpreters (European Union of the Deaf, 2018; World Federation of the Deaf, 2018; DeMeulder, 2021). Many proposed solutions have been developed by hearing researchers who have little if any at all connection with the

signing culture (Erard, 2017). Given this, the poor quality of sign language display is one of the major reasons for the skepticism or even hostility against avatar technology (Sayers et al., 2021).

The direct involvement of deaf users in the development and evaluation of signing avatars is imperative in order to eliminate skepticism, raise trust, and move forward with technologies acceptable by their consumers. A paradigm of constructive cooperation between researchers and the deaf community is the EASIER project⁴, where user driven design and technology development have already started producing results. One of the major goals of the project is the direct involvement of SL users at every stage of development of the project avatar. As developers wished to consider every parameter of SL articulation including affect and prosody, it was necessary to develop a steady communication channel with a wide public of SL users, who act as evaluators and provide guidance throughout research steps. To this end, we have developed a questionnaire-based methodology, which enables researchers to reach signers of different SL communities on-line and collect their preferences on various aspects of research work. In the next section, we report on the methodology behind the application of the EASIER evaluation framework for end-user guidance in signing avatar development.

Methodological framework of the on-line survey application

To identify how human signers perceive and evaluate the performance of an avatar's synthetic signing we have developed a shell environment which allows creation of on-line questionnaires to be addressed to various signer groups and question different aspects of synthetic signing performance. The first questionnaire application supported a survey on the preferences of signers regarding the display of affect, hand movement, hand and finger configuration accuracy in isolated signs and in fingerspelling, and smoothness of transition in short phrases, as performed by two synthetic signing engines. The questionnaire was distributed among members of the Greek Sign Language (GSL) community (the questionnaire of the survey can be found at: <https://sign.ilsp.gr/slt-survey/>).

Next, we present the survey questionnaire's structure along with the decisions and the methodological approaches adopted toward common and uncommon biases that occurred at every phase of its development.

Starting from the design of the shell methodological framework, we tried to create an as possibly unbiased environment which would maintain user-friendly characteristics. To do so we considered various parameters regarding the overall layout of the questionnaire, how the

⁴ <https://www.project-easier.eu/>

questionnaire would be distributed and the profile of the participants it would be addressed to, along with our need to regularly address end-users while proceeding with different stages of technological development. With all these parameters in mind, decisions on questionnaire content led to focused, short lasting questionnaire implementations. From a statistical point of view, an exhaustive questionnaire in terms of categories and items would provide a global view of the users' preferences. However, it would demand that the participants devote a significant amount of time and effort to complete it, which would turn its application prohibitive.

One of our main concerns was to balance between a reasonable questionnaire duration (maximum 20 min) that would not cause discomfort or fatigue to the participants, and adequate content to provide clear data on the intended head-to-head comparisons of synthetic avatar signing instantiations for which we needed user feedback. By setting up a viable and reproducible on-line survey we opted to engage into a steady dialogue with signers' communities with respect to various enhancements in the signing avatar technology.

For this first survey the questionnaire was divided in two parts. In the first part, the selection of signs that were generated by the two avatar engines was weighed upon the criterion of complexity with respect to handshape formations, manual movements and basic affect features. In the second part, isolated signs are mixed with short phrases focusing on motion of the upper body, the head, the eyes, and the mouth.

The exploited avatar technologies

The survey involved a head-to-head comparison between two signing avatars, Françoise, and Paula, representing the two most advanced avatar engines with generative capacity currently available. The two avatars use different strategies to create a display of signing performance including the manual element formations as well as the non-manual expressive markers. However, they both use the same original reference recording of productions in GSL that are part of the POLYTROPON bilingual lexical database (Efthimiou et al., 2016), currently comprising ~8,600 entries for the pair GSL-Modern Greek. All lemmas in the database are enriched with phonetic transcriptions according to the HamNoSys coding system (Hanke, 2004). This transcription enables synthetic signing productions *via* animation through an avatar character (Efthimiou et al., 2019).

Françoise was developed by the University of East Anglia (UEA) (<http://vh.cmp.uea.ac.uk>) in the framework of the Dicta-Sign project. She is animated by a SiGML script deriving from the HamNoSys notation strings which is stored on the SiS-Builder server (<http://sign.ilsp.gr/sisbuilder/index.php>), a tool that enables the creation and interrelation of SL lexicon entries with the HamNoSys features necessary to drive their synthetic signing and animation (Goulas et al., 2010).

Paula is an avatar developed at DePaul University (<http://asl.cs.depaul.edu/>) following 20 years of research on synthetic signing animation. Paula's design aims to produce linguistically correct signed outcomes that are convincingly natural in appearance and easy to understand.

The survey entails an experiment based on a selection of signs from the POLYTROPON GSL dataset; the original GSL videos and their accompanying HamNoSys notations were used for the signed productions on both avatar engines.

Greek sign language demographics

In Efthimiou et al. (2014) the demographic data about the Greek Sign Language (GSL) are presented as follows: “GSL is used by 1% of the 10 million people of the overall Greek population (Facts about Greek Deaf Population 2002), with several thousands of native and non-native signers. In 2000 GSL was approved by the Ministry of Education, as the official language for schooling of deaf persons, following recognition of GSL by the Greek Parliament as one of the official national languages of the Greek State (Legislative Act 2817/2000).”

Outreach of the first on-line survey

Over the past 10 years the Greek Sign Language and the Greek Deaf Community have been at the epicenter of research performed in various academic fields, such as psychology, education and educational policies, sociology, and linguistics. As a result, the Greek Deaf Community have been targeted as potential participants in multiple surveys, which for various reasons—that are not in the scope of the present article—were not adapted properly in the three-dimensional modality (i.e., questions presented in written text) nor were their outcomes fully disseminated for the wellbeing of the Community. Hence, the members of the Deaf Community have become reticent in participating in such surveys; being aware of this fact allowed us not only to adapt our survey in a fully accessible manner but also circulate our questionnaire *via* collaborating institutions from within the Deaf Community.

The sample of the population to which the survey was conducted, consisted of Greek Sign Language signers who can be broadly categorized in two groups; “L1 signers” including deaf, hard of hearing or hearing signers that acquired GSL from their immediate family environment from early childhood, and “L2 signers” including deaf, hard of hearing or hearing signers that acquired GSL *via* educational procedures (Costello et al., 2006).

The research team, composed of deaf and hearing GSL experts, has collaborated over the years with a significant number of partners with expertise in GSL and Deaf Studies; for the purpose of the present study, we addressed the on-line

questionnaire to the following bodies with the request to forward it to their GSL signer members:

1. The Department of Special Education of the University of Thessaly, with deaf and hearing GSL expert staff.
2. The Deaf Association of North Greece (Ένωση Κωφών Βορείου Ελλάδος).
3. The Association of Greek Sign Language Teachers (<https://sdideng.gr/>), having as members all deaf teachers active in Greece.

Due to GDPR issues and research ethics guidelines and regulations, responding to the questionnaire was anonymous. Moreover, we restricted personal information to a minimum set of metadata concerning demographic information on gender, age group, education level and GSL manner of acquisition (L1 vs. L2) that were necessary for the analysis of the results. No additional information regarding hearing condition, social and educational status was requested.

During the 3-week period that the questionnaire was circulated among GSL signers before we calculated the results, 91 distinct IP addresses were identified as having visited the questionnaire. By the end of the 3-week period, only 32 out of the 91—one third—had completed the questionnaire, while the rest had stopped responding at various parts of it. One can think of a variety of factors for the questionnaire drop out, including the extremely hot weather conditions during the period, little interest for the topic of the survey or interest decreasing gradually over time, or a combination of all. This indirect feedback will allow us to revisit the overall design of the survey framework including the timing of future distributions.

Regarding the analysis of results, we have taken into account only those questionnaires for which the participants provided data for all questionnaire pages, and hence were considered completed.

Questionnaire instructions display

One of our main concerns was to provide a survey shell fully adapted to the three-dimensional language modality. Considering that language is the principal factor for interaction, we ensured that all questionnaire parts and items could be accessible with the use of sign language only. Hence, in every stage of the questionnaire participants were provided with instructions as to what they were expected to evaluate and how they could interact with the questionnaire environment (i) *via* GSL videos recorded by a L1 signer of GSL, (ii) *via* written text (Greek) available to be viewed if selected, in a text box below each instructive video, and (iii) *via* screen capture videos demonstrating the requested action by the user.

An introductory video presented the scope of the questionnaire, the identity of the research team and a brief description of the EASIER project.

All questionnaire instructions were recorded in the premises of ATHENA/ILSP, in a recording studio that qualifies the highest recording standards. For these recordings, a L1 GSL signer presenting instructions was captured by a High Definition (HD) camera. All videos were rendered with the Advanced Video Coding (AVC) H.264 (MPEG-4) format.

The total duration of the two parts of the survey, including the duration of the video instructions, did not exceed 20 min.

Structure of the first on-line survey questionnaire

The first on-line questionnaire was structured in two parts. In Part A, both avatars were presented to participants on the same screen in a head-to-head manner, while in Part B, participants viewed one avatar at a time. Special care was taken so that in those questionnaire pages where both avatars appeared, these were presented in similar body and face dimensions subject to display settings of the two distinct avatar technologies and against a similar background to minimize bias.

The adopted structure allowed for the collection of information on a variety of aspects, which are analyzed in Section Results.

The linguistic content of the questionnaire was distributed in the two parts as follows:

Part A: participants were presented with both avatars head-to-head, and they were asked to evaluate:

- (i) Avatar expressivity *via* inspection of still images of avatar face pairs, while depicting the emotions of JOY, FEAR, ANGER, SURPRISE, and SORROW (Figure 1),
- (ii) Pairs of avatar productions of the following signs: TOMORROW, LAKE, INTERPRETER, TRAIN, MILK, HISTORY, BREAK (Figure 2),
- (iii) Pairs of avatar performance while fingerspelling the proper names MANOS, NASOS, MARIA,
- (iv) Pairs of avatar productions of a set of four short phrases including the previously evaluated signs along with other signs not yet viewed by participants.

In total, 19 signing instances, grouped into 4 categories, were examined as illustrated in the first two columns of Table 1. The presentation order of the two avatars was randomized to avoid bias in the responses. Viewers had to mark their preference, but also rank the performance of the two avatars.

Viewer preference between the two avatar displays in Part A was indicated by color code (red frame in Figure 1), while viewers had to also provide their ranking regarding signing performance of both avatars (boxes marked in yellow in Figure 1) before they could move forward to the next page of the questionnaire. This was a checkpoint of special interest since it could reveal further information regarding viewer attitude than the indication of preference only.

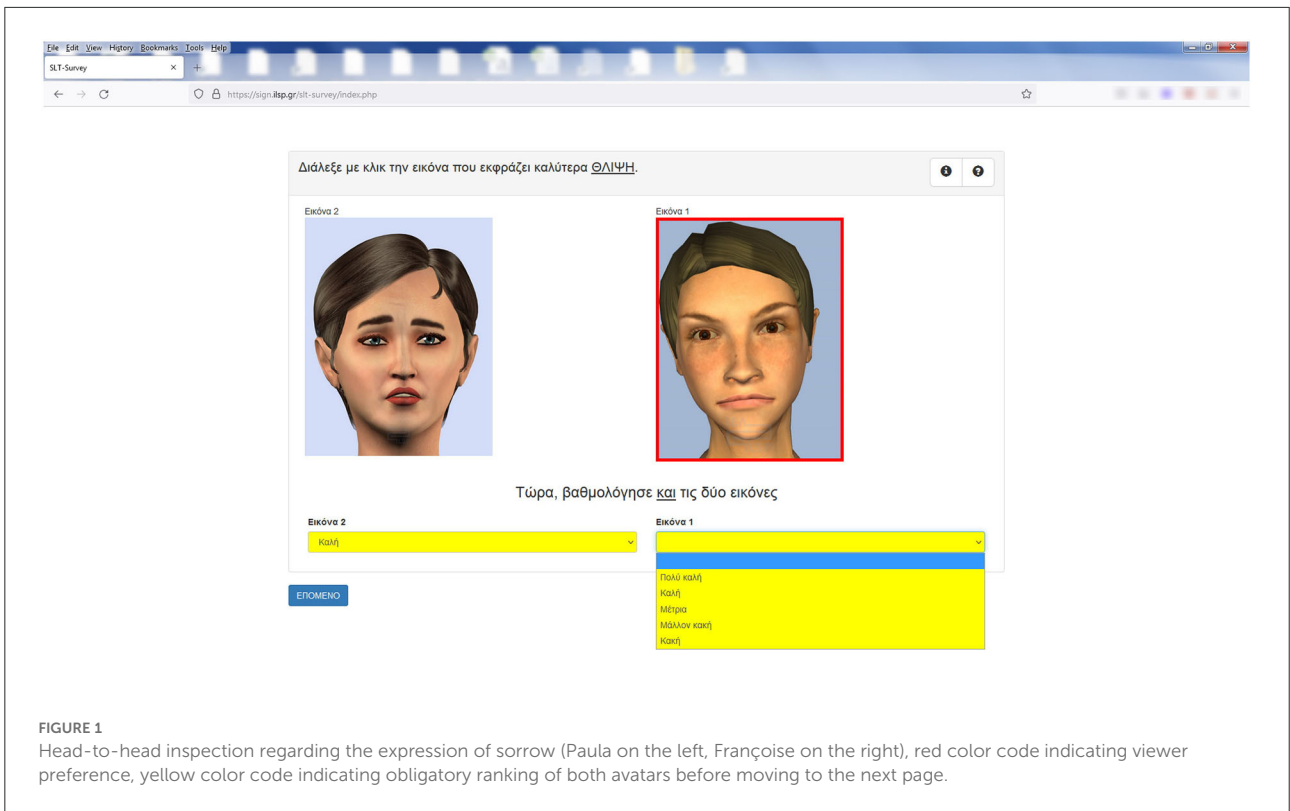


FIGURE 1 Head-to-head inspection regarding the expression of sorrow (Paula on the left, Françoise on the right), red color code indicating viewer preference, yellow color code indicating obligatory ranking of both avatars before moving to the next page.

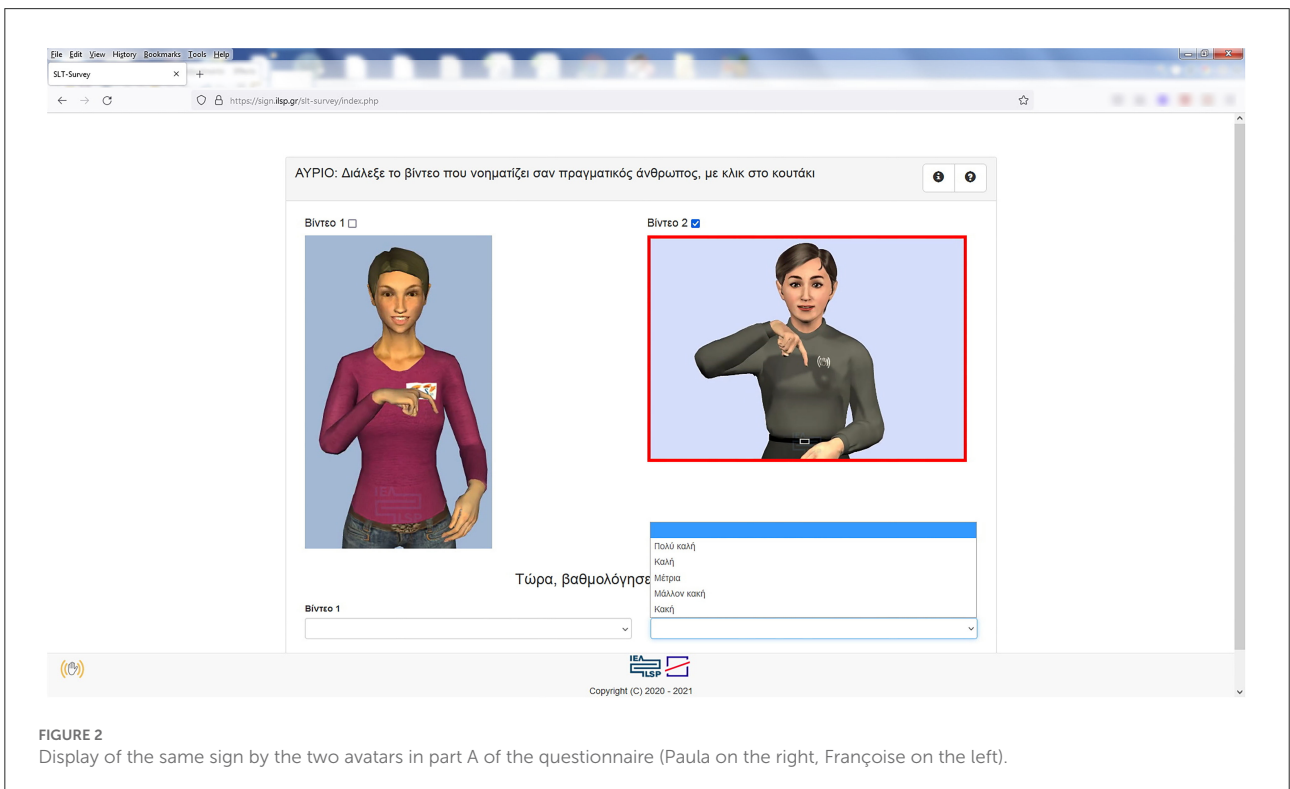


FIGURE 2 Display of the same sign by the two avatars in part A of the questionnaire (Paula on the right, Françoise on the left).

TABLE 1 Analytic overview of median, minimum, maximum, and quartile values for all signing occurrences presented to the participants for both avatars in Part A.

Groups	Signing occurrences	FRANÇOISE					PAULA					
		Min	1st quartile	Median	3rd quartile	Max	Min	1st quartile	Median	3rd quartile	Max	
Emotions	Joy	1	3	3.5	4	5	1	3	4	4	5	
	Sorrow	1	2	3	4	5	2	4	4	5	5	
	Surprise	1	3	3	4	5	2	3	4	4	5	
	Fear	1	1	2	2	5	1	4	4.5	5	5	
	Anger	2	4	4	5	5	1	3	4	4	5	
Lemmas	Tomorrow	1	2	3	3	5	2	4	4	4	5	
	Lake	1	2	3	3	4	2	4	4	5	5	
	Interpreter	1	2	3	4	4	3	4	5	5	5	
	Train	1	3	3	4	5	1	3	4	4	5	
	Milk	2	3	3	4	5	2	4	4	5	5	
	History	1	3	3	4	4	3	4	5	5	5	
	Break	2	3	3	4	5	2	4	4	5	5	
	Fing/ling	Manos	2	3	3.5	4	5	2	4	4	5	5
		Nasos	2	3	4	4	5	3	3	4	5	5
Maria		2	3	4	4	5	3	4	4	5	5	
Phrases	Phrase 1	1	2	3	3	5	3	4	4	5	5	
	Phrase 2	2	3	4	4	5	2	3	3.5	4	5	
	Phrase 3	2	3	3	4	5	3	4	4	5	5	
	Phrase 4	1	3	3	4	5	3	4	4	5	5	
Total		1	3	3	4	5	1	4	4	5	5	

The wording of the tasks posed to the participants was a subject of study and discussion among the research team. It was decided to avoid questions of the form “*which avatar do you prefer?*” since they could possibly lead to judgements regarding the external appearance of the avatars, while the aim was to gather data about specific avatar performance features. This led to the decision that the most suitable task formulation for this specific questionnaire would not include the term “avatar” and would neither be phrased as a question but would focus on the signing/emotion production instead. Hence, for the three stimuli categories which were presented head-to-head including lemma, fingerspelling and phrase productions, the viewers were given the task “*choose the video in which the signing performance is similar to human signing by clicking on the box.*” For the stimuli category of still avatar images expressing emotions the viewer task was “*choose the image that expresses [EMOTION_TYPE] best by clicking on it.*”

In Part A after choosing the closer to human avatar performance, participants also provided a ranking of the performance of both avatars. Ranking of avatar performance was based on a Likert scale from 1 to 5 corresponding to Bad/Rather Bad/Average/Good/Very Good. In order to guarantee that participants go through all protocol steps they had to complete all ratings on a screen before being able to move to the next screen (yellow color indication in Figure 1).

In Part B participants were presented with one avatar at a time. Each avatar performed a set of signs and short phrases. In this part each of the two avatars displayed different content (Figure 3). The aim of this part was to lead viewers to focus on specific features of interest in each avatar performance directly linked to the underlying driving technology. Tasks included rating each avatar separately in respect to:

- (i) overall hand motion performance,
- (ii) overall body motion performance,
- (iii) head and eyes movement,
- (iv) mouth movement.

Thus, in this part participants responded only by providing a ranking of each avatar with respect to tasks such as: “*how do you evaluate the hand motion performance?*” Ranking avatar performance was based on the same Likert scale as in Part A, ranging from 1 to 5 and corresponding to Bad/Rather Bad/Average/Good/Very Good.

The survey was accessed by GSL signers via the online questionnaire available in <https://sign.ilsp.gr/slt-survey/>, in which participants were able to watch avatar productions in the form of embedded videos. Regarding software technologies, the questionnaire was created using the open-source Cascading Style Sheets of the Bootstrap Framework. Bootstrap is a framework that allows the creation of responsive, mobile-first

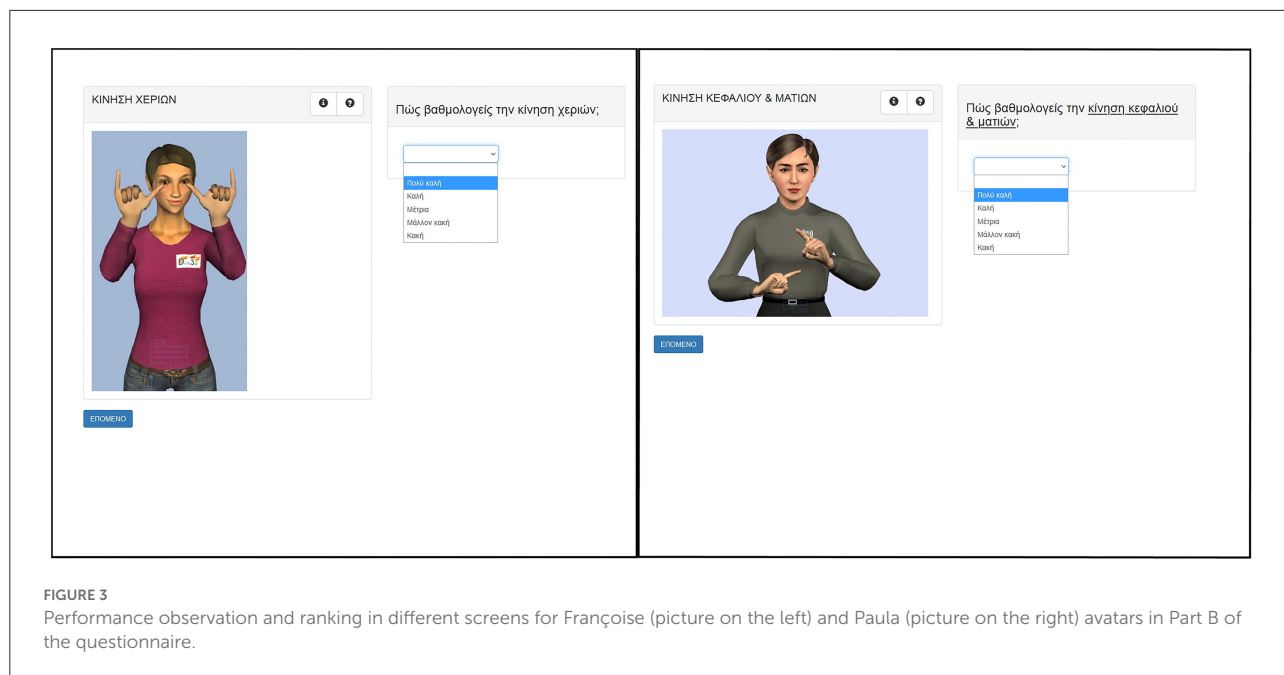


FIGURE 3 Performance observation and ranking in different screens for Françoise (picture on the left) and Paula (picture on the right) avatars in Part B of the questionnaire.

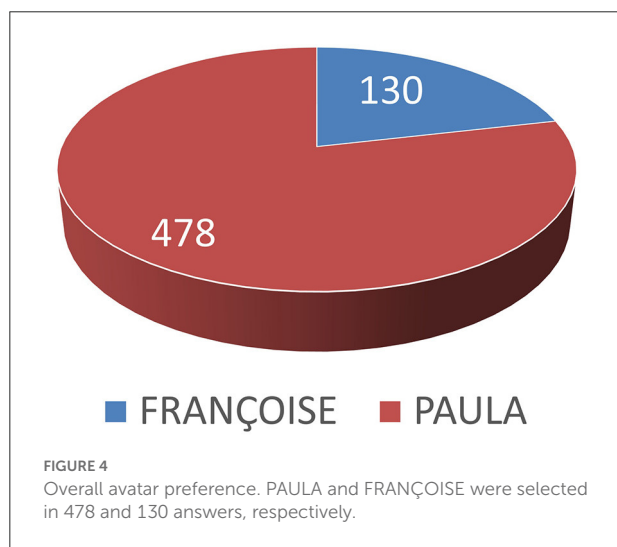


FIGURE 4 Overall avatar preference. PAULA and FRANÇOISE were selected in 478 and 130 answers, respectively.

web applications. Thus, web applications created by Bootstrap Framework can be executed by most desktop as well as mobile browsers. However, due to the considerable number of images and videos in the application, participants were encouraged to use Firefox or Chrome for optimum performance. The user interface was created using HTML5 and JavaScript (jQuery). The database in which participants' answers are stored is MySQL. Php is used to store the data in the database.

Results

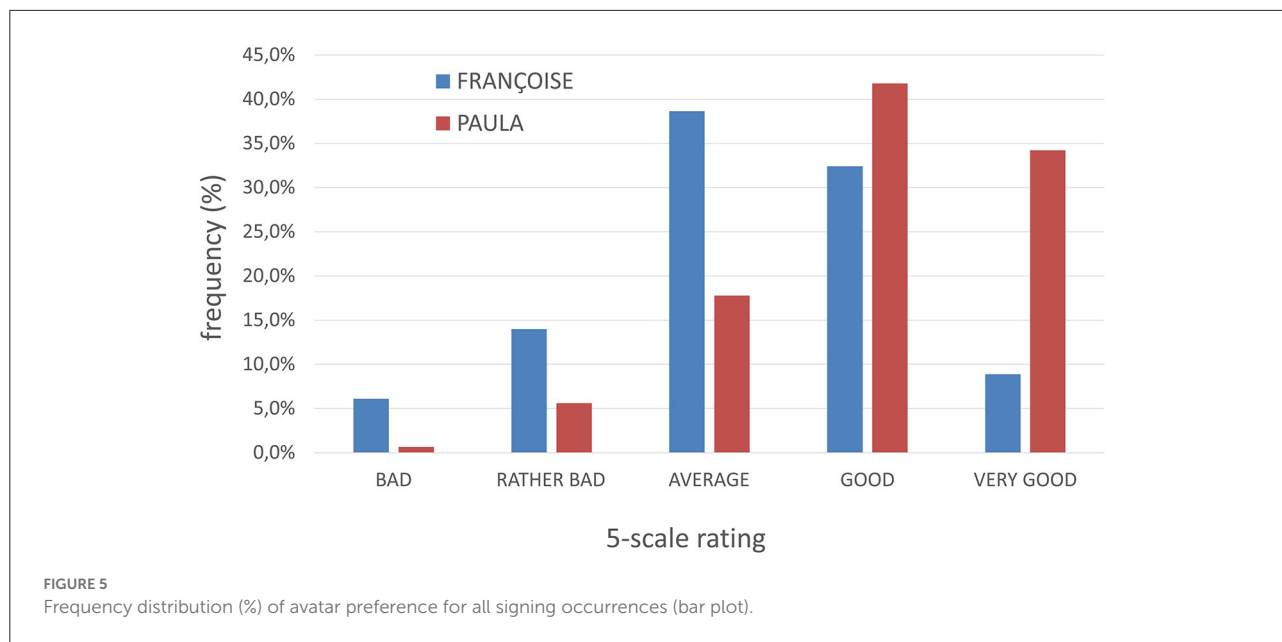
Participants' profile

Ninety-one GSL Signers Participated in the Survey in Total, but Only 32 of Them Completed the Questionnaire. Thus, Only the Data From Those 32 Have Been Accounted for in the Result Section.

According to the metadata information provided by the participants at the beginning of the questionnaire, 17 among the 32 participants identified themselves as L1 signers, having learned GSL in their immediate family environment and 15 as L2 signers, having learned GSL in an educational setting later in life.

Among the 32 participants, 21 were female and 11 were male. They were all adults with an age distribution ranging from 18 to 61 years. Due to the restricted number of participants, it was not attempted to perform statistical tests with respect to the metadata parameters. However, within the scope of the overall survey framework, we envisage that future distribution of the questionnaire will be able to provide us with data that will satisfy the statistical requirements for such statistical analysis.

As mentioned above, participation in the survey was voluntary. This fact implies that we did not select a random sample of the targeted population. Thus, no statistical tests were carried out to show significant statistical comparisons between the two avatars, but only descriptive statistics are presented.



Part A results

In the context of part A, each participant was asked to select the avatar that signed closer to a human in each of the 19 occurrences, grouped within the questionnaire in the following four categories: i. Emotions: 5 images, ii. Lemmas: 7 videos, iii. Fingerspelling: 3 videos and iv. Phrases: 4 videos. Out of the obtained 608 answers (32 participants* 19 sign occurrences), PAULA was selected in 478, while FRANÇOISE in 130 (see Figure 4).

In addition to the closed question, participants were asked to rate the performance of each avatar in each signing occurrence in a 5-scale rating (Bad/Rather Bad/Average/Good/Very Good). To get an overview of the obtained data from this 5-scale rating, the frequency distribution in percentages is illustrated in the following bar plot (Figure 5).

By observing the graph above, we conclude that

- a) The mode (i.e., the most frequent response) for the totality of the signing occurrences is “Average” for FRANÇOISE and “Good” for PAULA.
- b) PAULA’s frequency distribution is more right skewed than FRANÇOISE’s one.

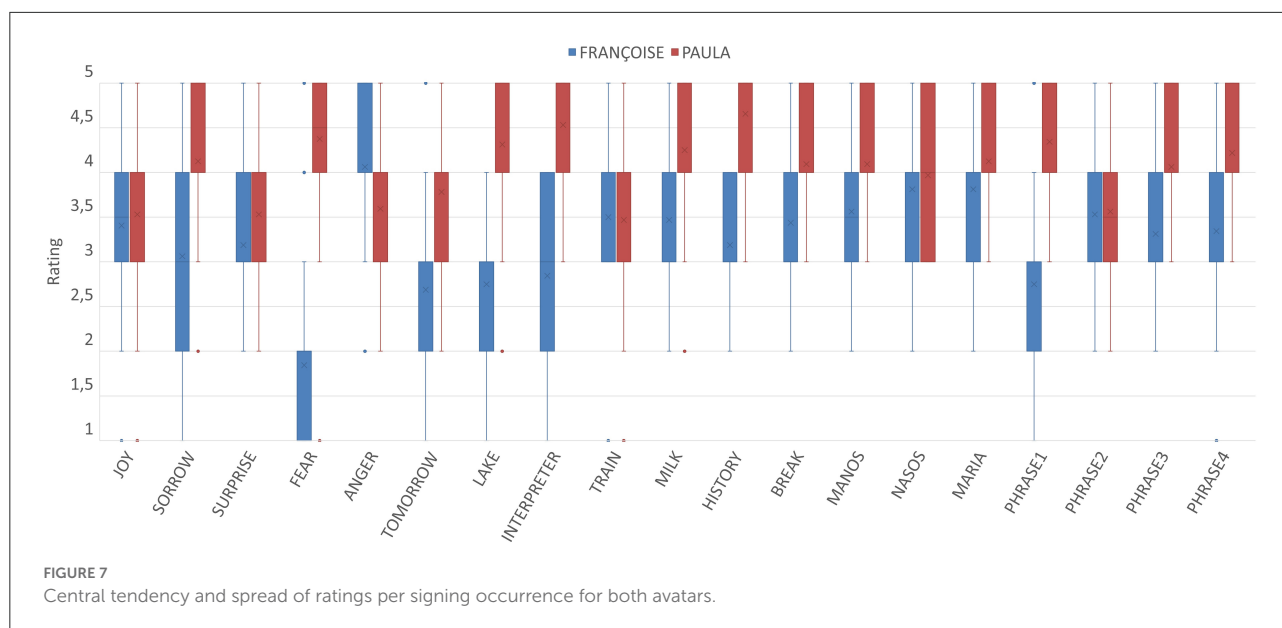
Both these findings are consistent with the participants’ judgment on the binary question which of the two avatars signs closer to a human; the participants preferred PAULA over FRANÇOISE.

To visualize the central tendency and the spread of the collected data (per signing occurrence and in total), we sorted the data in an increasing order, calculated the minimum,



maximum, median (i.e., midpoint of the distribution of the ordered dataset) and quartiles values for all signing occurrences presented to the participants within the four groups (see left column) of content for both avatars (see Table 1), and generated the boxplots depicted in Figures 6, 7.

In almost all cases, the median value equals either to the 1st quartile and thus the median line in a boxplot coincides with the lower boundary of the box/rectangle, or to the 3rd quartile and thus the median line coincides with the upper boundary. There are four cases in which the median values do not seem meaningful (e.g., 3.5 or 4.5) but these results are due to the even number of the evaluators (i.e., 32). The obtained data from



the 5-scale rating are ordinal, hence the mean value cannot be computed. However, a light “X” in each box has been added, arbitrarily, at the position of the mean value, with the purpose of helping readers who are used to numerical data, to interpret our data. In Figure 6, the boxplots present the central tendency and spread for all signing occurrences for each avatar (see last row of Table 1).

Regarding ratings for all signing occurrences, the median for FRANÇOISE is 3 (Average) while for PAULA is 4 (Good). Based on the blue box’s height and position, we observe that the answers for FRANÇOISE’s performance are concentrated at 3 and 4 (Average – Good) while PAULA’s ones (see red box’s height and position) are concentrated at 4 and 5 (Good – Very Good). Moreover, the blue whiskers visualize the spread of the answers given for FRANÇOISE on both sides on the main “lobe,” show that rates for FRANÇOISE range from 2 (Rather Bad) to 5 (Very Good) and 1s (Bad) are considered outliers (see blue dot). Respectively, the answer’s for PAULA’s performance range up to 3 (Average), considering 1s and 2s outliers (red dots).

The following figure (Figure 7) is more explanatory of the performance of each avatar as it presents the central tendency and spread for each signing occurrence the participants were exposed to and illustrates their preferences.

We observe that in most signing occurrences PAULA obtained a higher rating than FRANÇOISE. The most interesting findings from this illustration concern the still images for “JOY” and “SURPRISE,” the lemma “TRAIN” and the utterance “PHRASE2.” For these signing occurrences the participants evaluated the performance of both avatars as similar, even though the median values for all these occurrences is 3 (Average) for FRANÇOISE and 4 (Good) for PAULA and is consistent with the general tendency as seen in the previous figure.

The first two occurrences are images expressing emotions, those of “JOY” and “SURPRISE.” The lemma “TRAIN” (means of transportation) received similar evaluation for both avatars. The signing occurrence “PHRASE2” is the phrase that signifies “*In the train there are many seats.*” Interestingly the still image for “SORROW” and the lemma “INTERPRETER” as performed by FRANÇOISE received responses that have significant variation; the ratings for FRANÇOISE are more disperse 2–4 (Rather Bad – Good), while the great majority of PAULA’s rating for the same signing occurrences are 4 or 5 (Good – Very Good). Contrary to almost all other cases in which the median line of a boxplot coincides with either the lower boundary of the box/rectangle (i.e., the 1st quartile equals to the median value), or the upper boundary (i.e., the 3rd quartile equals to the median value), in these two cases the 1st quartile, median and 3rd quartile equal to 2, 3, and 4, respectively (see the horizontal line in the middle of the respective box).

Finally, the image expressing the emotion of “ANGER” as expressed by FRANÇOISE received higher ratings than the one expressed by PAULA.

Part B results

In the second part of the questionnaire, the participant task was to provide ratings on the individual performance of each avatar with respect to four movement parameters: i. hand(s) movement, ii. body movement with emphasis on the movement of the shoulders, iii. head and eyes movement and iv. mouth movement. For each of these parameters the participants watched different video compilations that consisted of two lemmas and two phrases. We selected videos in which each avatar was performing best regarding these parameters. Hence,

the created compilations included different content regarding the two avatars. Participants provided their ratings of the signing performance of each avatar on the same 5-scale (Bad/Rather Bad/Average/Good/Very Good).

To get an overview of the obtained data from this 5-scale rating, the frequency distribution in percentages is illustrated in the following bar plot (Figure 8).

By observing the graphs in Figure 8, we conclude that the mode (i.e., the most frequent response) for both avatars, FRANÇOISE and PAULA, is “Good.” This is an interesting finding as it indicates that the overall impression of the signing performance of both avatars is equally satisfying the participants.

To visualize the central tendency and the spread of the collected data, we generated the boxplots in Figure 9. The boxplots present the central tendency and spread overall, for all movement parameters for both avatars. It is obvious that the values of 1st and 3rd quartiles are equal to 3 (Average) and 4 (Good), respectively.

Moreover, the whiskers show that the ratings vary from 2 (Rather Bad) to 5 (Very Good) while 1s (Bad) could be considered outliers. In addition, we mention that median values for both FRANÇOISE and PAULA are equal to 4 (Good).

Figure 10 illustrates participants’ ratings for both avatars in the four parameters. We observe that in three out of four movement parameters FRANÇOISE obtained a stable rating, between 3 and 4 (Average – Good). PAULA’s ratings on the other hand varied more.

More specifically, with respect to “HAND(S) MOVEMENT,” PAULA performed better than FRANÇOISE. The ratings for “BODY MOVEMENT” are concentrated at 3 (Average) and 4 (Good) for both avatars. However, further analysis shows that the median value for FRANÇOISE coincides with the lower boundary of the rectangle, in this case 3 (Average), and the median for PAULA coincides with the upper boundary, in this case 4 (Good). Regarding “HEAD & EYES MOVEMENT” the median values are 4 (Good) for both avatars. However, we observe that for PAULA there are also higher ratings ranging up to 5 (Very Good). The “MOUTH MOVEMENT” is the only parameter in which FRANÇOISE obtained higher ratings than PAULA. The median for FRANÇOISE is 4 (Good) and 3 (Average) for PAULA.

Discussion

Statistical results interpretation

Regarding the first part (Part A) of the survey and the head-to-head presentation of the two avatars, for which participants were asked to choose the avatar that had a signing performance closer to the performance of a human, results showed that Paula was the avatar of preference.

Out of the total 608 signing occurrences (19 stimuli of images and videos multiplied by 32 participants), Paula was

chosen in 428 of them. Moreover, for each head-to-head instantiation, participants ranked the signing performance of each avatar; the statistical analysis showed that the most frequent response for the totality of the signing occurrences for PAULA is “Good” and for FRANÇOISE is “Average” (Figure 5). This finding is consistent with the obtained results from the head-to-head viewing task for deciding “*which of the two avatars signs closer to a human*” for which the participants expressed a preference for PAULA over FRANÇOISE.

The analysis of per-signing occurrence results showed that in most cases PAULA was the one that participants rated as closer to human.

Even though a larger amount of data is necessary in order to safely draw conclusions, we here attempt to interpret the results for these occurrences that stand out of the general tendency which favors PAULA’s signing over FRANÇOISE’s one.

Still images expressing emotions

“JOY” and “SURPRISE”: These images were rated in a similar way for both avatars. This finding allows us to assume that the expression in both avatars is equally satisfying for the viewers.

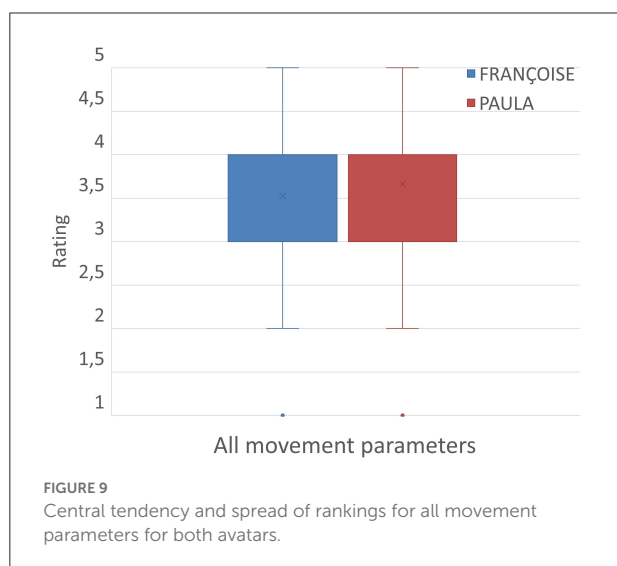
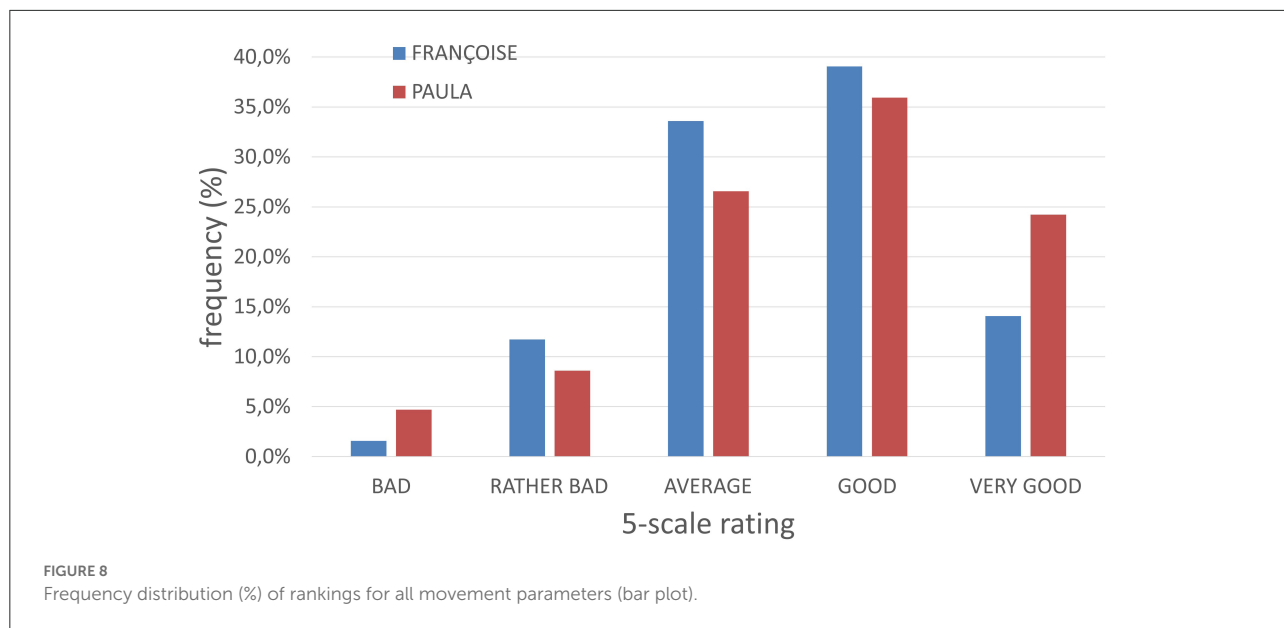
“ANGER”: This is the only signing occurrence -still image- for which FRANÇOISE gets higher ratings. Although a lot of research work still needs to be carried out in the domain of embedding emotion expression in synthetic signing, analysis of the expressive means of FRANÇOISE will provide significant insights as regards a complex set of implementation parameters, starting from the facial characteristics of FRANÇOISE in this emotion and the reasons why they were perceived as more convincing in comparison to those of PAULA’s.

“SORROW”: The image of the emotion as performed by FRANÇOISE received responses that presented significant variation ranging from Rather Bad to Good. PAULA on the other hand received more stable ratings (Good – Very Good). Participants in general preferred PAULA over FRANÇOISE, while the variation in FRANÇOISE’s ratings presents a further interesting point for research.

Overall, findings generate interesting research questions with respect to the facial articulators (i.e., eyes, eyebrows, mouth etc.) that participate in the creation of emotion expression in signing avatars. Additionally, we need to investigate intensity and width of facial features and the way they combine in the expression of various emotions. Answering such questions will allow us to incorporate signers’ feedback in signing display technology.

Fingerspelling

For all three fingerspelling videos the participants expressed an explicit preference for PAULA over FRANÇOISE. Further testing is needed to validate this finding with more complex



strings of fingerspelling in isolated strings as well as within linguistic context.

Isolated lemmas

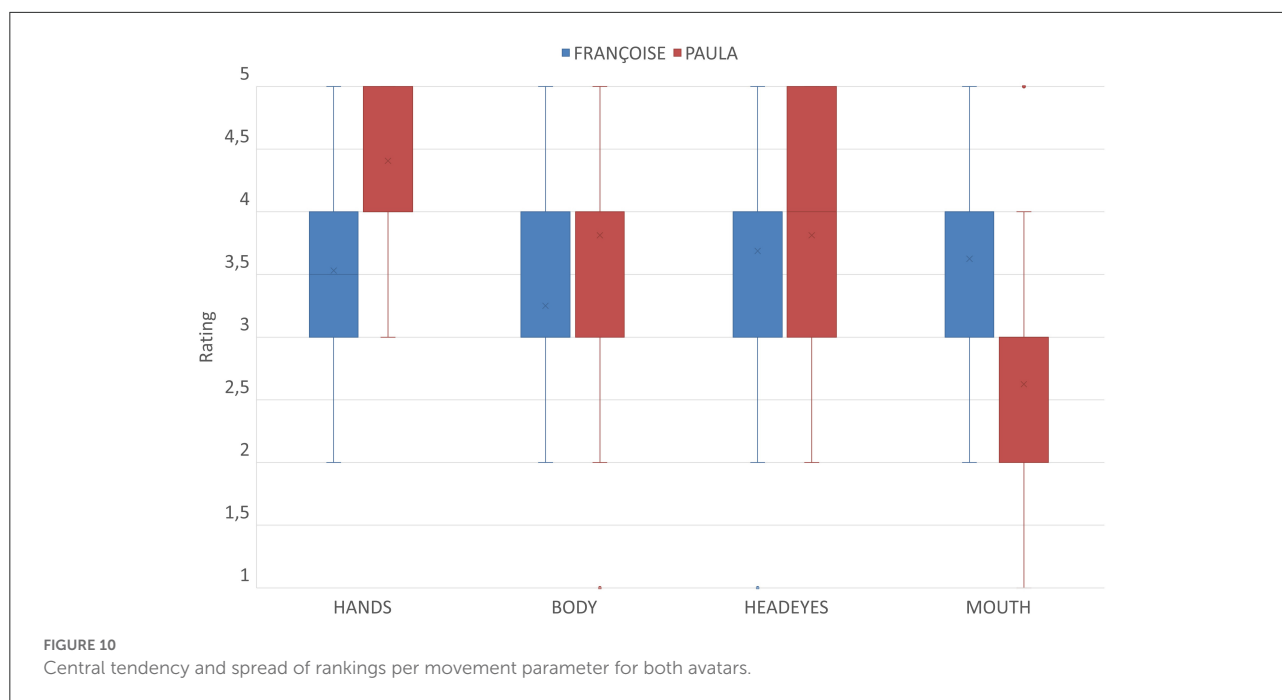
“INTERPRETER”: This GSL lemma as performed by FRANÇOISE received responses that had significant variation ranging from Rather Bad to Good. PAULA’s ratings for the same lemma varied from Good to Very Good. Like with the emotion image for SORROW, the significant variation in the participants’ responses indicates that the performance of this lemma needs to be reevaluated.

“TRAIN”: The lemma TRAIN was chosen in this questionnaire for the complexity of the hand movement it involves. Our goal was to receive feedback on the articulation capacity of the two avatar engines regarding performance of this specific sign, the production of which involves technically difficult movement and requires coordination of both hands. According to the participants’ judgment and their ratings, this lemma was equally well-performed by the two avatars (Average – Good). This finding becomes significant when combined with the respective findings from “PHRASE2” (see below).

Short phrases

“PHRASE2”: The second GSL phrase in the questionnaire roughly corresponds to the English translation “*There are many seats in the train.*” This phrase -similar to all others used in this questionnaire- is a small phrase containing some basic components of GSL phrase formation. However, a certain complexity level is noticed, as it only contains lemmas that are performed with both hands (two handed signs). The ratings of the performance of this phrase for both avatars were similar, and they ranged from Average to Good. The most interesting finding is that the phrase, which in purpose contained the lemma “TRAIN,” was rated in a similar way as the lemma “TRAIN” in the isolated lemmas section (see above). Initially, this finding allows us to presume an overall consistency on the participants’ ratings. However, to safely claim the validity of this finding, we need to extend testing in the future to a larger pool of stimuli that will involve rating of individual lemmas in isolation and in context.

In the second part of the survey (Part B), each avatar was individually rated for its signing performance with respect



to a compilation of signing occurrences consisting of isolated lemmas and phrases. The overall inspection of the collected data for Part B attests that both avatars performed equally well. An investigation of their performance with respect to the four movement parameters that were evaluated (hand movement, body movement, head and eye movement, mouth movement) led to the findings in Figure 10. PAULA received higher rankings for hand movement and eyes movement, while FRANÇOISE was preferred over PAULA for her mouth movement. Both avatars were equally evaluated with respect to their body movement. These are important findings that need to be investigated in more signing occurrences, within context as well as in isolated instantiations.

Our data in terms of number of participants is sufficient for an initial descriptive analysis as the one performed above. However, in order to further investigate the participants' choices and their respective ratings with respect to gender, age and SL manner of acquisition (L1 vs. L2), we need to extend our survey aiming at a broader randomly selected pool of participants.

Conclusion

The here reported findings from an on-line survey provided significant feedback not only with respect to the targeted aspects of avatar performance, but also regarding the structure of the follow-up surveys, currently under preparation, to address different SLs in the framework of the steady signer consultation strategy on avatar development as implemented in the EASIER project.

The first implementation of the on-line survey has demonstrated its effectiveness in achieving the human-in-the-loop factor in the development of signing avatar technology. Based on the analysis of the collected data, it also proved to make use of a methodologically sound environment in respect to both survey structuring and the display of survey material.

Among the most noticeable findings that allow us to presume an overall consistency on the participants' ratings is the fact that the latter rated similarly lemmas in their individual appearance and within context. Noticeable variation in ratings of a single avatar should be further investigated with respect to factors like age, gender, L1 vs. L2 status and educational level. However, overall good signing performance seems to be equally perceived by all signers.

Future research

The aim being to involve signers in signing avatar technology, we have developed a methodological framework which makes use of a shell environment that can take the form of on-line questionnaires of varying content. Planned accommodation of content from different SLs on various aspects of SL articulation will provide guidance to avatar technology development work, based on experience gained from the first application of the proposed survey methodology as reported here incorporating content from GSL.

The research team's goal is to open a steady communication line with signers in Europe, which will enable the active involvement of deaf European citizens in signing avatar technological enhancement.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

A-LD, KV, and AV contributed to the design of the evaluation protocol and the definition of the tested linguistic content. VP contributed to the statistical analysis of the results of the study. TG developed the on-line questionnaire used in the study and contributed to the creation of the displayed GSL content. S-EF and EE contributed in the design of the evaluation protocol and the development of the user interface. All authors contributed to the article and approved the submitted version.

Funding

The work presented here was supported by the EASIER (Intelligent Automatic Sign Language Translation) Project. EASIER has received funding from the European Union's Horizon 2020 Research and Innovation Programme, Grant Agreement no. 101016982.

References

- Adamo-Villani, N., Popescu, V., and Lestina, J. (2013). A non-expert-user interface for posing signing avatars. *Disabil. Rehabil. Assist. Technol.* 8, 238–248. doi: 10.3109/17483107.2012.704655
- Branson, J., and Miller, D. (1998). Nationalism and the linguistic rights of Deaf communities: linguistic imperialism and the recognition and development of sign languages. *J. Socioling.* 2, 3–34.
- Brun, R., Turki, A., and Laville, A. (2016). "A 3D application to familiarize children with sign language and assess the potential of avatars and motion capture for learning movement," in *Proceedings of the 3rd International Symposium on Movement and Computing* (New York, NY: Association for Computing Machinery), 1–2. doi: 10.1145/2948910.2948917
- Costello, B., Fernandez, F., and Landa, A. (2006). "The non-(existent) native signer: sign language research in a small deaf population," in *Theoretical Issues in Sign Language Research (TISLR) 9 Conference*, eds B. Costello, J. Fernández, and A. Landa (Floridianopolis: Editora Arara Azul).
- De Meulder, M., Krausneker, V., Turner, G., and Conama, J. B. (2019). "Sign language communities," in *The Palgrave Handbook of Minority Languages and Communities*, eds G. Hogan-Burn, and B. O'Rourke (London: Palgrave Macmillan), 207–232. doi: 10.1057/978-1-137-54066-9_8
- DeMeulder, M. (2021). "Is "good enough" good enough? Ethical and responsible development of sign language technologies," in *The Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL)*, ed D. Shterionov (Association for Machine Translation in the Americas), 33–42.
- Efthimiou, E., Dimou, A. -L., Fotinea, S. -E., Goulas, T., and Pissaris, M. (2014). "SiS-builder: a tool to support sign synthesis," in *Proceedings of the 2nd ENABLE International Conference on Using New Technologies for Inclusive Learning*, eds E. Shirley, S. Ball, and A. Starcic (York: Jisc TechDis), 26–36.
- Efthimiou, E., Fotinea, S.-E., Goulas, T., Vacalopoulou, A., Vasilaki, K., and Dimou, A.-L. (2019). Sign language technologies and the critical role of SL resources in view of future internet accessibility services. *Technologies* 7, 18. doi: 10.3390/technologies7010018
- Efthimiou, E., Fotinea, S. -E., Dimou, A. -L., Goulas, T., Karioris, P., Vasilaki, K., et al. (2016). "From a sign lexical database to an SL golden corpus - the POLYTROPON SL resource," in *Proceedings of the 7th Workshop on the Representation and Processing of Sign Languages: Corpus Mining, Satellite Workshop of the LREC 2016 Conference* (Portoroz: European Language Resources Association), 63–68.
- Erard, M. (2017). *Why Sign-Language Gloves Don't Help Deaf People*. *The Atlantic*. Available online at: <https://www.theatlantic.com/technology/archive/2017/11/why-signlanguage-gloves-dont-help-deaf-people/545441/> (accessed November 09, 2017).
- European Union of the Deaf (2018). *Accessibility of Information and Communication*. Available online at: <https://www.eud.eu/about-us/eud-position-paper/accessibility-information-and-communication/> (Accessed October 26, 2018).
- Goulas, T., Fotinea, S. -E., Efthimiou, E., and Pissaris, M. (2010). "SiSbuilder: A sign synthesis support tool," in *Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies (CSLT), Satellite Workshop of the LREC 2010 Conference* (Valetta: European Language Resources Association), 102–105.
- Hanke, T. (2004). "HamNoSys-representing sign language data in language resources and language processing contexts," in *Proceedings of the Workshop on Representation and Processing of Sign Language, Satellite Workshop of the LREC 2004 Conference* (Lisbon), 1–6. Available online at: <https://pdfs.semanticscholar.org/a716def562bccacd264385e54f3befdd8c5e91.pdf> (accessed October 16, 2020).
- Jennings, V., Elliott, R., Kennaway, R., and Glauert, J. (2010). "Requirements for a signing avatar" in *The Proceedings of the Workshop on Corpora and Sign Language Technologies (CSLT), Satellite Workshop of the LREC 2010 Conference* (Valetta: European Language Resources Association), 33–136.
- Kipp, K., Nguyen, Q., Heloir, A., and Matthes, S. (2011). "Assessing the deaf user perspective on sign language avatars," in *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '11)* (New York, NY: Association for Computing Machinery), 107–114. doi: 10.1145/2049536.2049557

Acknowledgments

Huge thanks go to John McDonald and the students in the School of Computing at DePaul University and Rosalee Wolfe of the Institute for Language and Speech Processing in ATHENA RC for their invaluable help in preparing the GSL content displayed by the Paula avatar for the purposes of this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Krausneker, V., and Schügerl, S. (2021). *Avatars and Sign Languages: Developing a Best Practice Protocol on Quality in Accessibility*. Récupéré sur University of Vienna. Available online at: <https://avatar-bestpractice.univie.ac.at/> (accessed March 06, 2022).
- Kusters, A., De Meulder, M., and O'Brien, D. (2017). "Innovations in deaf studies: critically mapping the field," in *Innovations in Deaf Studies. The Role of Deaf Scholars*, eds A. Kusters, M. De Meulder, and D. O'Brien (Oxford: Oxford University Press), 1–53.
- McDonald, J., Wolfe, R., Moncrief, R., and Baowidan, S. (2016). "A computational model of role shift to support the synthesis of signed language," in *The Proceedings of the 12th Theoretical Issues in Sign Language Research (TISLR)* (Melbourne, VIC), 4–7.
- Pauser, S. (2019). *Prototypentest SiMAX im Rahmen des Innovationsschecks*. Available online at: https://www.equalizent.com/fileadmin/user_upload/News/2019_04_Avatar_Projektbericht.pdf (accessed March 16, 2020).
- Radford, A., Luke, M., and Chintala, S. (2016). "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proceedings of the International Conference on Learning Representations ICLR 2016* (Puerto Rico). doi: 10.48550/arXiv.1511.06434
- Sáfár, E., and Glauert, J. (2012). "Computer modelling," in *Sign Language. An International Handbook*, eds R. Pfau, M. Steinbach, and B. Woll (Berlin: De Gruyter Mouton), 1075–1101. doi: 10.1515/9783110261325.1075
- Sayers, D., Sousa-Silva, R., Höhn, S., Ahmedi, L., Allkivi-Metsoja, K., Anastasiou, D., et al. (2021). "The dawn of the human-machine era: A forecast of new and emerging language technologies," in *Report for EU COST Action CA19102. Language in The Human-Machine Era*. doi: 10.17011/jyx/reports/20210518/1
- Schödl, A., Szeliski, R., Salesin, D. H., and Essa, I. (2000). "Video textures," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY: ACM Press/Addison-Wesley Publishing Co), 489–498. doi: 10.1145/344779.345012
- Stoll, S., Camgöz, N. C., Hadfield, S., and Bowden, R. (2018). "Sign language production using neural machine translation and generative adversarial networks," in *The Proceedings of the 29th British Machine Vision Conference* (New Castle: BMVA Press).
- Wolfe, R., McDonald, J., Efthimiou, E., Fontinea, E. -S., Picron, F., Van Landuyt, D., et al. (2021). "The myth of signing avatars (long paper)," in *The Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL)*, ed D. Shterionov (Association for Machine Translation in the Americas), 33–42.
- Wolfe, R., McDonald, J., and Schnepf, J. C. (2011). "An avatar to depict sign language: building from reusable hand animation," in *Proceedings of the International Workshop on Sign Language Translation and Avatar Technology Workshop* (Berlin). Available online at: https://scholarworks.bgsu.edu/vcte_pub/13
- World Federation of the Deaf (2018). *WFD and WASLI Statement of Use of Signing Avatars*. Available online at: <https://wf-deaf.org/news/resources/wfd-wasli-statement-use-signing-avatars/> (accessed March 14, 2018).