

Authors

Alfarabi Imashev and Tohid Alizadeh
e-mail: alfarabi.imashev@nu.edu.kz; tohid.alizadeh@nu.edu.kz

Affiliations

National Laboratory Astana, Nazarbayev University; Robotics Department, School of Engineering and Digital Sciences, Nazarbayev University



NAZARBAYEV UNIVERSITY
NATIONAL LABORATORY ASTANA



Objective

The most extensive studies of sign language in Kazakhstan, pertaining to international academic standards, started around 10 years ago. Native deaf signers in Kazakhstan commonly face limited educational options, thereby leading usually to poorer reading skills. Often, deaf signers may identify letters and read words, although they might fail to understand the primary concept and need fragmentation into a series of simpler concepts for better understanding.

Thus, signing avatars possess the capability of interpreting online statements, film subtitles, or YouTube videos, thereby enhancing accessibility and fostering communication between deaf and hearing individuals, as well as between humans and avatars. An equally significant problem is the creation of a tool designed to allow deaf signers to assess the performance, aesthetics, and credibility of signing avatars without dependence on written questionnaires, especially for underresourced sign languages.

This paper delineates the iterative design of the Kazakh-Russian Sign Language interpreting avatar, ongoing refinements to the evaluation apparatus, and a comparative examination of such an instrument with an alternative assessment approach aimed at achieving the same purpose.

Introduction

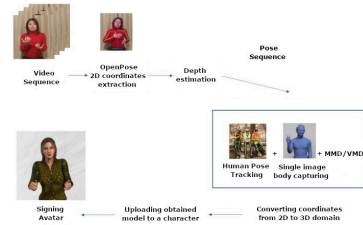
- Signing avatars may help native deaf signers and improve information accessibility, according to the research (Kipp et al., 2011a; Adamo-Villani et al., 2016; Bouzid et al., 2016). However, more recent research has found it challenging to create user-friendly, grammatically accurate, and deaf-friendly avatars (Bragg et al., 2019). Deaf people sometimes criticize sign language interpreting systems for lacking cultural relevance, accuracy, and the capacity to transmit subtle expressions and feelings (Wolfe et al., 2022). Therefore, developing proper sign language generation is challenging.
- Many deaf signers, despite usually being bilingual or multilingual (Gibet et al., 2011), identify sign language as their first language. A fundamental challenge to this undertaking is the absence of agreement over the complicated nature of doing evaluative studies with the deaf population. A significant problem is the implementation of deaf-friendly survey instruments in order to enhance researchers' comprehension of the responses from native deaf subjects.
- Conventional written questionnaires, often used for surveys with hearing respondents, are sometimes unsuitable for user assessment research, including with the deaf, so using a questionnaire in a written format, as is often performed, is inappropriate (Farwell, 1976). This strategy may compromise the validity of the feedback acquired for various evident reasons (Bosch-Baliarda et al., 2019). Hence, interaction with native deaf participants should be conducted in their first language (Gibet et al., 2011; Bosch-Baliarda et al., 2019). User assessment research, with evaluations that involve deaf participants, often finds conventional textual questionnaires unsuitable because they do not effectively capture the nuances of communication in sign language and may lead to misunderstandings.
- Another major challenge is the development of deaf-friendly survey metrics, which aims to enhance researchers' understanding of responses from deaf participants. The participation of the deaf population in the research process is also crucial in the assessment of sign language systems, since their perspectives might mitigate obstacles in experimental design and ensure that the research methods are inclusive and accurately reflect their communication needs.

Experimental Setup

Avatar Design



The evolution of K-RSL avatars



The pipeline of the updated approach used for versions 3.1 and 3.2

Evaluation tools iteratively included into the proposed questionnaire

User study #1 (Online user study)

- Five generic questions
- Four sections taken from Godspeed

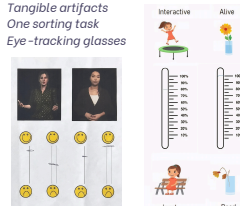
User study #2 (1st in-person user study)

- Thirteen generic questions translated into K-RSL as questions and provided as short videos
- Four sections taken from the Godspeed Questionnaire with all scales translated into K-RSL as questions, provided as short videos (10 questions)
- Two sorting tasks
- Eye-tracking glasses



User study #3 (2nd in-person user study)

- Funometer scales for emotional state
- Eleven generic questions translated into K-RSL as questions and provided as videos
- Four sections taken from the Godspeed Questionnaire with all scales translated into K-RSL as questions, provided as short videos
- Thermo-meter-like 100-point scales
- Tangible artifacts
- One sorting task
- Eye-tracking glasses

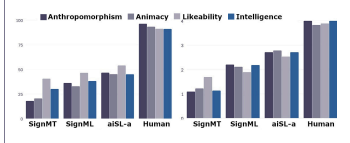


Comparison with EASIER

Advantages of EASIER

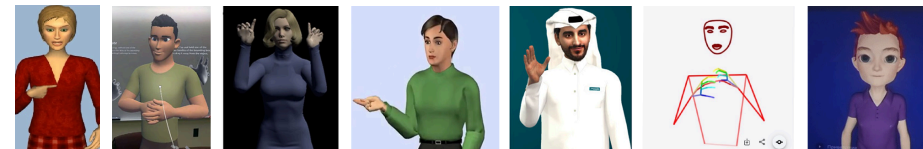
- Created as an online tool
- Shorter duration: 20 minutes vs. 40 minutes for the proposed questionnaire
- Evaluation of fingerspelling
- Individual selected lemmas also included
- Short signing sequences that contain the selected lemmas included
- Static visuals of emotions included
- A more complex understanding of the potential demerits of avatars to be evaluated
- Involvement of L1 and L2 participants

A difference revealed after the use of the 100-point scale.



Authors suggest utilizing 100-point scales, particularly when the participant cohort in the user study is small.

Existing alternatives and feedback from participants



The most well-known, reputable, and novel signing avatar alternatives currently available, from left to right: SIGML; the 3D avatar taken from (Yang et al., 2022) developed for a college mixed-reality application; the avatar from (Nguyen et al., 2021); Paula; BuHamad; SignMT; and Project Adaptis.

According to participants (local deaf signers) who had interacted with some of these avatars, despite overall positive perception, they also provided their opinions about demerits:

SIGML Anna: Most participants were dissatisfied with the agent's lack of eye contact, tendency to glance down, and lateral head motions. Such actions might be perceived as a reluctance to engage in dialogue or even as rude behavior in deaf culture.

SignMT: entangled fingers, which resulted in difficulties in differentiating them in motion. Furthermore, disproportionately big thumbs and limited mouth opening affected their ability to interpret the meaning of the performed signs.

Project Adaptis: limited vocabulary; a disproportionately large head, which makes the performance of some signs incorrect.

Results

- The iterative approach enabled the formulation, development, and refinement of a practical and quite effective questionnaire to evaluate signing avatars by native deaf signers.
- User studies revealed that native deaf signers still prefer interactions with human agents and are less optimistic about signing avatars than interpreters but still welcome this idea for future use.
- Several versions of local sign language avatars have been tested.
- More natural signings led participants to focus on the signing avatars' heads (faces), similar to interaction with a human agent.

Conclusions and Future Work

- A separate section for exterior attributes might be required.
- Perceived SL Expertise section formulation.
- Providing extra artifacts or concepts in the central sections of 100-point scales for clarity.
- A Text-To-Gloss LLM model trained to convert written text to gloss sequences with proper word order is needed.
- The participation of a deaf interpreter as a member of the research team throughout the preparatory stages and the user studies themselves will also be advantageous.
- It is still an ongoing process that requires testing other measures, techniques, and a greater participant cohort.

Related Literature

- Athanasia-Lida Dimou, Vassilis Papavassiliou, John McDonald, Theodore Goulas, Kyriaki Vasilaki, Anna Vaccapolou, Stavroula-Evita Fotinea, Eleni Efthimiou, and Rosalee Wolfe. 2022b. Signing avatar performance evaluation within easier project. In Proceedings of the 7th international workshop on sign language translation and avatar technology: The junction of the visual and the textual: Challenges and perspectives, pages 39–44.
- Frankie Picon, Davy Van Landuyt, Rehana Omardeen, Eleni Efthimiou, Rosalee Wolfe, Stavroula-Evita Fotinea, Theodore Goulas, Christian Tismer, Maria Kopf, and Thomas Hanke. 2024. The easier mobile application and avatar end-user evaluation methodology. In Proceedings of the LREC-COLING 2024 11th Workshop on the Representation and Processing of Sign Languages: Evaluation of Sign Language Resources, pages 276–281.
- Alfarabi Imashev, Nurziya Orabayeva, Gulmira Baizhanova, and Anara Sandygulova. 2025a. Assessment of comparative evaluation techniques for signing agents: a study with deaf adults. Journal on Multimodal User Interfaces, 19(1):1–19.