Assessing the Deaf User Perspective on Sign Language Avatars

Michael Kipp, Quan Nguyen, Alexis Heloir
DFKI
Embodied Agents Research Group
Saarbrücken, Germany
firstname.lastname@dfki.de

Silke Matthes
University of Hamburg
Institute of German Sign Language and
Communication of the Deaf
Hamburg, Germany
silke.matthes@sign-lang.uni-
hamburg.de

ABSTRACT

Signing avatars have the potential to become a useful and even cost-effective method to make written content more accessible for Deaf people. However, avatar research is characterized by the fact that most researchers are not members of the Deaf community, and that Deaf people as potential users have little or no knowledge about avatars. Therefore, we suggest two well-known methods, focus groups and online studies, as a two-way information exchange between research and the Deaf community. Our aim was to assess signing avatar acceptability, shortcomings of current avatars and potential use cases. We conducted two focus group interviews (N=8) and, to quantify important issues, created an accessible online user study (N=317). This paper deals with both the methodology used and the elicited opinions and criticism. While we found a positive baseline response to the idea of signing avatars, we also show that there is a statistically significant increase in positive opinion caused by participating in the studies. We argue that inclusion of Deaf people on many levels will foster acceptance as well as provide important feedback regarding key aspects of avatar technology that need to be improved.

Categories and Subject Descriptors

I.2.7 [**Artificial Intelligence**]: Natural Language Processing—*language generation, machine translation*; K.4.2 [**Computers and Society**]: Social Issues—*assistive technologies for persons with disabilities*

General Terms

Acceptance, Experimentation, Measurement

Keywords

German Sign Language, Sign Language Synthesis, Accessibility Technology for Deaf People

1. INTRODUCTION

Sign language avatars or *signing avatars* could be a useful tool for the many deaf[1] people who use sign language as their preferred language. For a deaf person, learning to read and write a spoken language, without auditory cues, is an inherently difficult task. Studies have shown that many deaf pupils leave school with significant reading/writing problems [6]. This implies that access to written content, e.g. in the internet, is limited for many deaf individuals. Therefore, an increased use of video-recorded human signers can be observed. However, video recordings imply considerable production cost, their content cannot be modified after production, and they cannot be anonymized with the face being a meaningful component of sign language. In contrast, when using signing avatars, i.e. virtual characters that perform sign language, one can change appearance (gender, clothes, lighting), they are inherently anonymous and the production of new content is potentially easy and cost-effective (no studio setup, no expert performer required, may even be created collaboratively) [8]. Most importantly, avatar animations can be dynamic, i.e. they can be computed and adjusted on-the-fly, allowing for the rendering of dynamic content (e.g. inserting locations, dates, times ...) and interactive behavior (question answering). In this paper, we focus on avatar technology that allows for this flexibility.

However, new technology always faces the question of *acceptability* in the targeted user group. Given a generally small proportion of Deaf people in research positions, signing avatars are almost exclusively developed by hearing researchers [2]. Deaf individuals may be skeptical about any technology invented by the hearing for historical reasons. However, the question of acceptance is essential for the success of a later implementation of such a technology, and therefore also crucial for governmental and other agencies when deciding on funding. While prior work has invested considerable effort in involving Deaf people in animation [7] and evaluation [16], a larger effort to clarify general acceptance is, to the best of our knowledge, still missing. Acceptance implies identifying potential negative sentiments or fears concerning this technology. Exposure to and assessment of current avatar technology is prerequisite for such an assessment, and there is the potential of eliciting new ideas

---

[1] We follow the convention of writing *Deaf* with a capitalized "D" to refer to members of the Deaf community who use sign language as their preferred language, whereas *deaf* refers to the audiological condition of not hearing [15].

for avatar applications. Ultimately, not only the assessment of acceptance but also the question of how to increase acceptance must be addressed. In the study reported here we found that the mere participation in our user studies increased acceptance to a measurable degree.

Signing avatars have been in the focus of research for over two decades [8, 12, 10]. Important goals of this research are methods for translating from spoken languages to sign languages, notation systems to describe sign language and animation methods to automatically create natural and comprehensible sign language movements with avatars. It is important to stress that automated animation is not nearly as natural as hand-animated movies (Toy Story or Avatar) or computer game animations. Therefore, while significant results have been achieved, the sentence-level comprehensibility of avatars remains relatively low, averaging around 60%, with a single result of 71% in a particular scenario [10]. Compared to spoken language processing research, the community is small and lacks the budget to create the same international networks that have fostered spoken language research.

To investigate the potentials of signing avatars for the internet, the German Federal Ministry of Labour and Social Affairs (Bundesministerium für Arbeit und Soziales, BMAS) commissioned us to investigate the technical feasibility of signing avatars for German sign language (DGS[2]) and the acceptance in the German Deaf community. In this paper, we focus on the acceptance aspect of this study. The major goals of this study were to include Deaf people early on to identify key aspects, send out a signal to the Deaf community that we intend a close cooperation and to develop methods for conducting studies with Deaf people.

We see the following contributions for the research community:

- A combination of methods for assessing signing avatar acceptability, identifying shortcomings of current avatars and eliciting ideas for possible applications. Our methods consist of pure sign language dual-moderator focus groups, complemented by an accessible internet questionnaire.

- Identifying problematic aspects of existing avatars from the perspective of German Deaf users. These comprise mostly *nonmanual* aspects, especially facial expression, mouthing and torso movement.

- Showing that, in Germany, Deaf people have a mildly positive "baseline" attitude towards avatars and, more importantly, that this positive attitude can be increased by participating in either focus group or online study.

In this paper, we first outline the background in opinion mining within the Deaf community, then we present our methods, including design and participation, and analysis, before we discuss results and limitations and conclude with a summary and future work.

2. BACKGROUND

In this section, we survey the potential methods for opinion mining and highlight work done in the context of the Deaf community. The research area of signing avatars is

---

[2]Deutsche Gebärdensprache

located on the crossroads of *language technology*, *linguistics* and *human-computer interaction* (HCI). In contrast to spoken language technology the targeted language community of Deaf people is underrepresented in active research [2]. This makes it especially important to closely include Deaf individuals at all stages of research. When considering methods from HCI, we distinguish methods depending on the phase in development (early stage vs. prototyping) and the qualitative/quantitative dimension (Fig. 1). In this paper, we focus on user studies at an early stage of development.
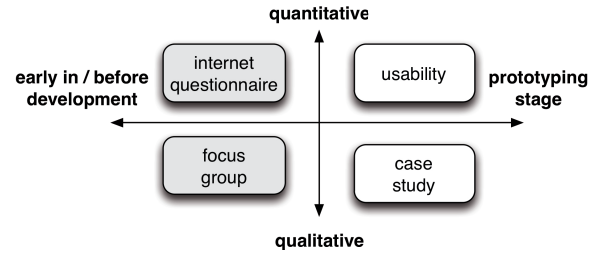

Figure 1: In the space of possible HCI techniques, focus groups and internet questionnaires combine the virtues of qualitative and quantitative studies.

Focus groups are a well-known HCI method to elicit empirical data and are applied in marketing, political campaigning and social sciences research. The goal is to elicit people's perceptions and attitudes about any particular product or concept early on in the design process before actual prototypes exist. Focus groups offer a context for comments, interaction, and exchange, thereby giving access to *in-depth* information about what issues participants consider important, what preferences they have and how they prioritize these. A focus group is a guided discussion with 3-10 participants led by a trained moderator where a preset agenda guides the discussion. In contrast, *single case studies* have been shown to be highly useful for eliciting not only usability issues but also socially-situated side conditions, for instance the concern of a blind person to be "marked out" as being blind by a speaking watch [17]. Focus groups with Deaf people were pioneered in 1999 in five groups with an experience hearing moderator and a sign language interpreter/assistant [1]. It was stressed that a shared mode of communication (sign language) in a non-threatening atmosphere needs to be established. Deaf focus groups in a technology context have been conducted for American Sign Language (ASL) in the *MobileASL* project [3] to elicit requirements and scenarios for video cell phone usage for Deaf people. They conducted a one-hour session with four participants where a hearing researcher was present mediated by a sign language interpreter. For the *ClassInFocus* project [4] about visual notifications a two-hour, loosely structured focus group plus prototype testing session was conducted with eight participants. Other researchers used one-on-one interviews with Deaf individuals to elicit user needs [5, 13]. Matthews et al. [13] interviewed 8 Deaf subjects about the design of a sound visualization system. The first part contained structured questions, the second part asked for open feedback on ten design sketches. Tran et al. [18] claim to be the first to design an internet questionnaire (N=148) specifically for Deaf people, testing the usability of MobileASL, and point
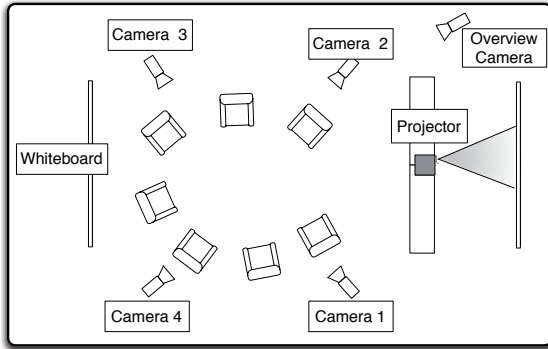
(a) Setup of the second focus group G2.

(b) Screenshot of the video used for the analysis.

Figure 2: Focus group setup and analysis video.

out the importance of offering both ASL and textual explanations, both for understandability and to show respect for Deaf culture.

Usability testing in the area of signing avatars is mostly concerned with the *comprehensibility* of avatar signing. To assess the comprehensibility of the sign language output produced by an avatar is not straightforward and no agreed-on methodology exists. [9] found that there is a low correlation between Deaf participants own judgement on how well s/he understood a sentence and the objective assessment of understanding. A number of in-depth comprehensibility tests have been done in the ViSiCAST/eSIGN projects [10, 16]. These tests usually featured a rather low number of participants. However, more recent evaluation studies use much higher quantities of participants and materials [7]. We present a method for comparing avatar performance directly with video-recorded human signers in [11] to complement the work presented here.

3. METHOD

Our approach is to combine in-depth discussion and the elicitation of ideas that is possible in focus groups with the quantitative strengths of online studies. We suggest the following improvements over prior work on focus groups with Deaf individuals: Ensure a pure sign language environment [9], rely on mostly visual materials (icons, images and video) [14], combine open discussion with structuring and voting, and complement focus groups with accessible internet studies [18] for the quantification of results.

3.1 Focus groups

The overall aim of the focus groups was to elicit opinions, criticism and priorities early in the project and to actively involve the Deaf community. We wanted participants to discuss signing avatars in a non-threatening environment, i.e. without hearing researchers in the room, and in their preferred mode of communication, i.e. sign language [1]. We decided for a *dual-moderator focus group*. The main moderator would guide the discussion, while an assistant would be able to prepare keyword cards for "voting" phases, where participants could indicate their priorities with stickers. The assistant would also act as a communication link with the (hearing) researchers for the case that clarification was needed. As a moderator we hired a well-connected Deaf

community member and sign language instructor, while for the assistant we hired a formerly deaf person, fluent in German Sign Language (DGS), with cochlear implant who could speak.

For technical reasons, participants had to sit in a circle, tables were removed to not obstruct the view. All sessions were videotaped for later analysis (all subjects signed an agreement to grant us scientific usage of the material). The videos were recorded by five cameras which had to be set up so as to minimize occlusion of participants by other participants (see Fig. 2(a)). For later analysis the videos had to be synchronized and cut together in one video (Fig. 2(b)). Synchronization is best done with a clear visual/auditory marking at the beginning (e.g. a person clapping).

We conducted two focus groups, G1 and G2, with 3 and 5 participants each. Each group took about four hours. Participant selection was done by the moderator and assistant according to the following criteria: Participants should be native signers and should consider themselves members of the Deaf community. For G1 we aimed for more computer literate, educated and open-minded members. For G2 we wanted a more representative sample of different education levels. Each of the 8 participants (6m, 2f), of age 25-50, was compensated with 30 Euro plus travel cost. Because of our selection criteria we had to include more remote cities for recruiting (up to 360 km in one case).

During the focus groups we used different media to stimulate discussion: a projector was used to show videos of existing avatars and still images e.g. to illustrate suggested applications, and a whiteboard was used to stick flash cards with keywords on it that could be used for voting (participants could put red dots on those flash cards they found most important). Accept from the keywords, written text was avoided throughout the session.

Each focus group was structured in cycles of information–discussion–voting. This way, we would guarantee sufficient background knowledge, active participation and a synthesis and quantification of results. Each focus group was first welcomed by the moderators and the scientific staff was introduced (outside the actual focus group room) to show that the researchers have genuine interest and respect and to create a pleasant atmosphere. Afterwards, subjects, moderator and assistant entered "their" focus group room. During the focus group, there were five blocks:

1. Introduction and initial questions The project was introduced by a sign language video and two initial questions were asked: "Do you think avatars are useful?" and "Do you think Deaf people would use avatars?" Answers were rated on a 5-point Likert scale from *not at all* to *absolutely* which were visually enhanced with smiley icons and color coding (red/yellow/green).
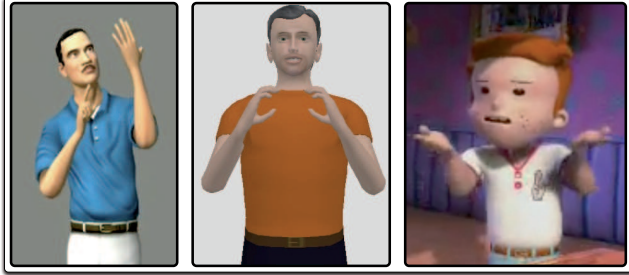

Figure 3: Three of the presented existing avatars (left to right): The Forest, Max and DeafWorld.

2. Avatar critique Videos of six existing signing avatars were shown (Fig. 3). Those avatars were The Forest[3] (ASL), Max[4] (DGS), DeafWorld[5] (International Sign), Sign4Me[6], the Grandpa Project[7] (BSL) and a Finnish signing boy[8]. This material was intended to give the participants a broad impression of what is possible today. We showed mainly avatars that could be used for automatic animation from a notation; such avatars typically look very robotic. Therefore, to allow participants to think about future possibilities we included the fully hand-animated *DeafWorld* clip. Since the range of existing avatars is quite limited we had to include avatars that used different sign languages (ASL, DGS, International Sign and others) and were communicating in different domains (government, poetry, festival invitation etc.). After each video, participants discussed and criticized the avatar. Keywords were taken and a round of voting on the relative importance of the keywords was conducted. Since this discussion turned out to be time-consuming, we reduced the amount of videos from six to four in group G2 because of the higher number of participants in this group.

3. General applications The participants saw images of several application scenarios, depicted on a photo montage, and discussed these. Our suggestions were: *at the doctor's*, *at the employment agency*, and *accommodation search*. Then, own suggestions were developed. Finally, votes for the most interesting applications were collected.

4. Internet applications This block was similar to the previous block, with the difference that applications specifically for the internet were presented, discussed and voted on. Our suggestions for internet applications were *online shopping*, *Deaf internet portals* and *forums*.

5. Final questions The participants were asked the first two initial questions once again. Additionally, we asked

---

[3] `www.youtube.com/watch?v=80L2XcOK8Jg`

[4] `www.einfach-teilhaben.de/DE/GBS/Home/Aktuelles/avatar_inhalt.html`

[5] `www.youtube.com/watch?v=QiY5LU-II6Q`

[6] `www.youtube.com/watch?v=-NfQGMrqEWY`

[7] `www.deadcreative.com/deadcreative/projects.asp`

[8] `www.youtube.com/watch?v=eI9DbYxdmzc`

the participants whether they thought that the government should invest money in this technology.

After these blocks, the participants left the room and the official wrap-up was done, including payment and personal discussion with the scientific staff. The analysis of the focus groups was conducted on the basis of the video recordings by a sign language interpreter who prepared an audio translation which was analyzed and summarized by a sign language researcher.
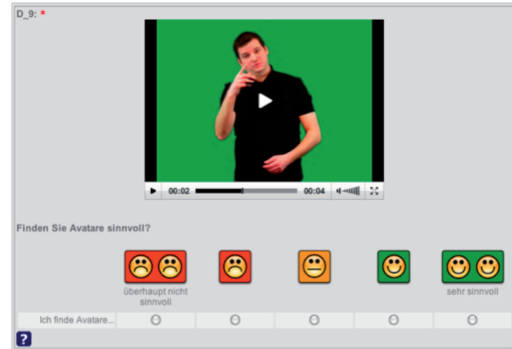

Figure 4: Our accessible online study provided DGS videos and used 5-point answer scales.

3.2 Online Study

To quantify several results from the focus groups we created an accessible internet user study (Fig. 4). It was based on the structure and results from the focus group. The online study was open to the general public but only advertised within the Deaf community via mailing lists, personal contacts, and the popular German Deaf web portal *www.taubenschlag.de*. For accessibility we provided DGS video explanations for all questions. Replies were collected on 5-point scales, visually enhanced by smileys and color-coding.
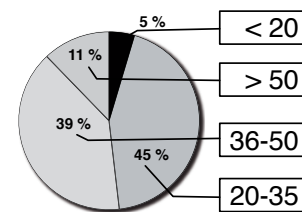

Figure 5: Age structure (online study)

The content of the survey was analogous to the blocks from the focus groups. In the avatar critique block, we only presented the three avatars shown in Fig. 3. Participants were asked how much they liked/disliked the following aspects of the avatars on a 5-point scale: comprehensibility, facial expression, naturalness, charisma, movements, mouthing, appearance, hand-shapes, clothing. In the other blocks we asked open questions, e.g. suggestions for applications after showing the same example applications as in the focus groups. Also, these questions were asked at beginning and end: "Do you think avatars are useful?" and "Do you think Deaf people would use avatars?". Additionally, we

asked "Do you think that the government should investigate money in this technology?" at the very end.

In total, 317 people completed the questionnaire. 42% were male, 58% female. There were more deaf (85 %) than hard of hearing (2%) or hearing (13%) participants. The age of the participants was quite balanced between the two largest age groups 20–35 and 36–50 (Fig. 5). On average, the participants had little or no experience with avatars. Fig. 6 shows the distribution of professional backgrounds. It is noteworthy that many people came from technical areas.
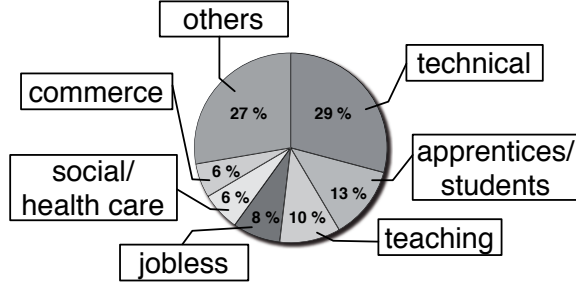

Figure 6: Professional background (online study)

4. ANALYSIS

In this section, we give a topic-oriented account of the results from focus groups and online study.

4.1 Existing Avatars

In focus groups and online study participants were shown 3-6 videos of existing avatars (Fig. 3) and were asked to criticize them. We first discuss our findings from the focus groups organized by topic.

Style and personality Most of the avatars presented were found to have hardly any emotional expression, therefore lacking charisma and naturalness. Their presence was described as stiff and sometimes robot-like. It was found that missing personality can easily be interpreted as cold or unfriendly. While the DeafWorld avatar, as a positive exception, was pointed out to have a very positive and pleasing presence, it was also noted that its cartoonish style may only be suitable for certain contexts.

Upper body movement and manual components No matter whether the language produced by the avatar was DGS or a foreign sign language, the focus of the participants' criticism was usually *not* on the manual component of the signs (i.e. the hands) but rather on the upper body movement as a whole. While the amount of movement varies between the avatars presented, the overall criticism was that it is not sufficient: except for hands and arms also a variety of head, shoulder, and torso movements are needed. For the torso, movements like hunching and twisting were mentioned to improve naturalness, as well as clear sideway rotations (e.g. for marking role shift). It was also found that there was a limited use of the signing space as the hands mainly show horizontal and vertical movements (e.g. pointing straight to the front instead of sidewards). In general, the participants wished for more smooth and relaxed movements of all parts of the upper body.

Facial expressions and mouth patterns While facial expressions are essential for sign languages to deliver

Important avatar aspects
Facial expression (7)
Natural movement (5)
Mouthing (4)
Emotions (4)
Body motion/posture (4)
Appearance (3)
Synchronisation of sign and mouthing (3)
Charisma (2)
Comprehensibility (2)

Table 1: Voting on most important avatar aspects (focus groups)

emotions as well as grammatical information, most of the avatars show very little. This was highly criticized by all participants, and for many avatars the face was described as stiff and emotionless. Specifically mentioned were missing variations in eyebrow, eyelid, as well as ocular movement. Permanent eye contact was regarded as unnatural and causing discomfort. The absence of mouth patterns, especially mouthings (i.e. mouth patterns derived from the spoken language), seemed to be one of the most disturbing factors for the participants since this is an important element of DGS. Besides more movement of cheeks and lips, teeth and tongue were said to be needed as a crucial element for understanding certain mouthings. While more facial and mouth movements seems to be essential, the participants stressed that exaggeration of movements should be avoided.

Movement synchronization The participants' feedback made very clear that the overall image is crucial for understanding the avatar's performance. For those avatars showing mouthing it was found that there was a mismatch between the duration of the signs and their corresponding mouthings. While commonly the signer's face is kept as a focus point, this mismatch provokes an disturbing oscillation of the observer's gaze between hands and face.

Technical remarks Good lighting and a clear contrast between the avatar's skin, clothes and the background is important for the perception of the signing avatar. Additional shadows were noted as favorable as they support a 3D effect. In order to meet individual needs speed and perspective should preferably be under user control.

Avatar appearance Of all avatars presented, the DeafWorld cartoon animation was ranked as the best one. However the participants emphasized the need of having different avatars for different domains: while a cartoonish child would be suitable for children and for entertainment, a more realistic adult avatar is recommended for the use in serious applications (e.g. politics).

Voting on important avatar aspects Table 1 shows what participants deemed most important for an avatar. It is striking is that many *nonmanual* components were mentioned whereas hand/arm movements were not explicitly mentioned. Also, the importance of realism and high-quality rendering should not be underestimated.

Online Study Results In the internet questionnaire we asked for ratings (between -2 to +2) for those aspects that were identified in the focus groups. For readability we merged the online study results for the avatars Forest and Max and compared this with the DeafWorld animation ratings. Fig. 7 clearly shows the large quality gap between a purely hand-

Application (general)	Application (internet)
Simple help/info dialogue (9)	Lexicon (7)
Train/airport (5)	News (4)
Fixed texts (4)	Education (3)
Forms (2)	Insurance (3)
Exam quest. (2)	Consumer protection (3)

Table 2: Voting on possible applications, only top 5 each (focus groups)

made animation like DeafWorld and the quality of automated avatars. We also compared this against the voting in the focus group and found a loose correspondence between aspects deemed important by participants and low values for the Forest/Max avatars. It was also apparent that nonmanual aspects are at least as important as the manual ones.
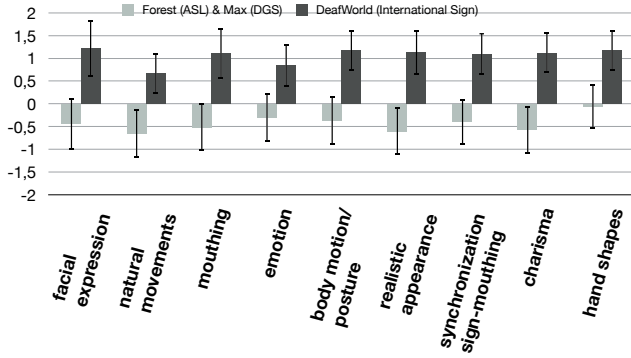

Figure 7: Ratings of aspects of the presented avatars (online study). Ratings for Forest and Max were averaged to juxtapose them against the fully hand-made DeafWorld animation.

4.2 Application Scenarios

In the focus groups, possible applications for avatars were mainly seen for one-way communication situations with less complex content. The participants could not envision dialogic interaction with an avatar. Many ideas emerged during discussion such as: (Online) translation services for simple sentences, static announcements (job offers, company newsletter, election campaigns) and static texts (legal texts, manuals), information usually communicated via speakers (train station, airport), daily news and news feeds, lexicons and dictionaries, museum guide.

When voting on the relative importance of applications there were quite concrete and technically realistic scenarios that won (Table 2), while avatars were not considered necessary for very trivial texts such as accommodation ads or restaurant menus.

The online study (Fig. 8) showed a much more diverse picture, probably due to the fact that the individual scenarios were not discussed with other Deaf people in terms of being technically realistic and actually relevant to everyday life. Also, more entertainment and leisure time applications came up. For internet applications the top applications were educational (17%), for social network websites (16%) and (public) administration pages (11%).
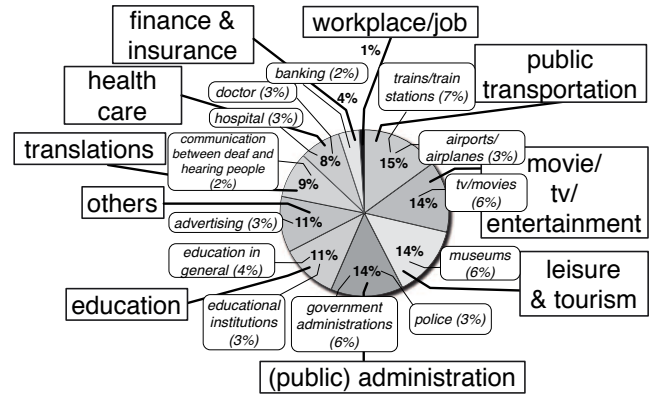

Figure 8: General applications (online study)

4.3 Risks and Potentials

In the focus groups, it was extremely important for all participants that avatars should not be seen as a replacement for human interpreters and that every Deaf person should always have the choice between the two. This is reflected in the online study where 25% of mentioned effects concerned job cuts for interpreters or for Deaf people. Another concern was the danger that using an avatar may lower the motivation for Deaf individuals to properly learn reading/writing. Maybe not surprisingly, the online study participants were much more concerned about technical feasibility (20%) and reliability (22%) of avatars since they had not discussed potential scenarios in depth. See Fig. 9 for the possible risks mentioned in the online study.
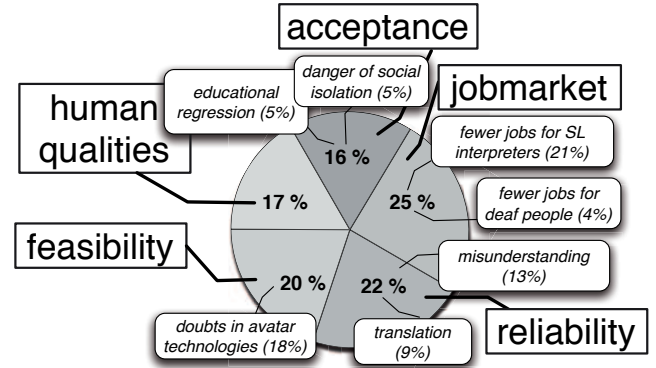

Figure 9: Fears of potential negative impacts in these areas (online study)

However, there were also a number of potentials seen in avatar technology. Focus group participants found it most important that avatars are *available anytime*, while interpreters are often hard to find. *Personalization* is possible, e.g. regarding appearance, speed, or perhaps even language output (sign language vs. sign supported spoken language). Avatars also allow for *anonymity* in the internet (e.g. for the discussion of controversial topics).

4.4 Acceptance

While all three participants of G1 already had a very positive attitude towards the use of avatars prior to the focus

group interview, they stated that their opinion was even higher after the discussion. For G2 the rating was initially neutral to slightly negative, but increased throughout the discussion to a slightly positive attitude. The participants regarded avatars as a good opportunity to provide full access to information in some domains. They assumed, however, that the acceptance across the Deaf community may vary depending on age, technological knowledge and sign language competence.

The general acceptance trend of the focus groups was confirmed in the online studies. The average value for the two questions at the beginning and the end ("Are avatars useful?" and "Would Deaf people use avatars?") on a scale of -2 (not at all) to $+2$ (very much so) was only slightly positive in the beginning ($M = 0.46; SD = 1.08$ and $M = 0.46; SD = 1.11$). In the end, the value increased significantly ($M = 0.74; SD = 1.12$ and $M = 0.71; SD = 1.03$). A paired t-test showed the that the increase was highly significant for both questions ($p < 0.001$ and $p < 0.001$). The question whether the government should invest in this technology was positive ($M = 0.77; SD = 0.77$).

5. DISCUSSION

5.1 Deaf User Opinions

The feedback of both the focus groups and the online study clearly shows that much improvement in the performance of sign language avatars is still needed. In the focus groups the criticism was not so much focussed on a single aspect than rather on the general appearance that was mostly described as stiff, emotionless and unnatural. Most of the criticism targeted nonmanual features, mainly facial expression and mouth patterns, but also movements of the head, shoulders and torso. The wish for naturalness and emotions also show that the general appearance should not be underestimated. On the contrary, manual components (i.e. the hands) were not in the participants' focus. This might be due to the fact that in former research comparatively much attention has been paid to improving the animation of the hands. However it also shows that other components are still underestimated in linguistic as well as technological research. The most positive votes in the focus groups as well as the online study were given to a fully handmade animation. This underlines the gap between avatar approaches which can be automated and handmade animations. However, it also shows that in principle animated characters can reach high levels of acceptance.

Application scenarios, as discussed in the focus groups, are mainly situated in the area of one-way communication situations. The participants can neither envision avatars in dialogic settings nor for very complex or emotional content. They also expressed worries regarding a potential replacement of human interpreters by avatars in these contexts. In the online study, an even higher emphasis was put on feared negative effects as job cuts when implementing avatars.

The general attitude towards avatars, as asked for in the beginning and at the end of the focus groups, was overall positive and increased throughout the sessions. In the online studies, this effect even reached high statistical significance. This underlines the potential of involving the Deaf community, not only for general assessments but also for increasing acceptance.

5.2 Method

The focus group interviews turned out to be an excellent method to elicit criticism, constructive suggestions and opinions of Deaf participants. Especially for a topic like avatars, where participants might not have a clear idea of the opportunities associated with this new technology, the focus group interviews allowed them to develop their criticism and suggestions throughout the session. We decided to establish a pure sign language environment. While this adds considerable overhead in terms of preparation (training the moderator) and analysis (video analysis) we believe that it creates an open atmosphere during the session, where new ideas can be discussed and clarification questions can be asked, and respect for the Deaf community is signaled by involving them in key positions of a scientific procedure. In terms of participants, we deem four participants the maximum number in terms of later analysis.

Thinking about the limitations of our approach, our results first of all apply to *German* Deaf users. However, we think that many aspects are valid for avatars of any sign language. In terms of participants, we would have liked to put more focus on specific groups (by age, by education, by gender) to be able to compare focus group results. Also, we would have liked a better male-female ratio. However, depending on the region, there may be very few Deaf people that meet all of the targeted criteria. We had a heterogenous mixture of avatar video material (different languages and domains). While this was positive in terms of conveying the breadth of research, it made comparison between avatars difficult and introduced order effects, e.g. seeing an avatar signing in the familiar DGS may have put more positive emphasis on it. Lastly, the topic of avatars often puts avatars in direct comparison with sign language interpreters. Since this is not a realistic near-future development the moderator should probably guide the participants more toward a comparison between avatars and videos (of signing humans).

In the focus groups, the participants' perception of the shown avatars were sometimes inaccurate or false. While a strength of focus groups lies in the possibility to correct this in the discourse, sometimes the contrary happens and the whole group is "infected" by an erroneous assumption, e.g. that mouthing was particularly good or bad. This is why complementary methods must be combined for an objective picture of Deaf people's opinions. Internet questionnaires can statistically validate the generality of some assumptions, e.g. which aspects are the most problematic with existing avatars. However, to focus more on avatar *performance* one has to measure the *comprehensibility* of an avatar in a test setup where a Deaf participant tries to understand an avatar's signing performance (see [11]).

6. CONCLUSIONS

We presented our approach for assessing the opinion of the German Deaf community about sign language avatars. Focus group interviews were complemented by a large internet study. We conducted the focus groups in a pure sign language environment and elicited valuable feedback on existing avatars and ideas on application scenarios. A significant increase in positive opinion in both focus groups and internet study showed that both methods help increase acceptance in the community. Therefore, including the Deaf community or Deaf individuals in this way works in two directions:

Not only does it benefit research, it also reverberates in the Deaf community through the dense, nationwide networks. We hypothesize that the positive influence on the opinion of a few will quickly spread throughout the community.

We found that nonmanual components were found to be at least as important as manual ones. This indicates that research needs to make a major shift toward new challenges in the nonmanual area. In terms of applications, Deaf people favored non-interactive, simple scenarios where avatars give information (train station, museums) or help in educational contexts (sign language lexicon, exam questions). However, also many other small everyday scenarios which may be made easier with an avatar were identified.

While focus groups and questionnaires give a good general impression of general shortcomings, the actual development of animated avatars needs a much deeper involvement of Deaf individuals. Comprehensibility studies allow to quantify the performance of an avatar by comparing its comprehensibility with that of a human signer on video. This can be measured by asking for a retelling or letting experts judge the participants' understanding. We explore these questions in a separate publication [11].

For the future we hope to conduct further focus group interviews on more specific topics, with better avatar materials or interactive mockup scenarios (like in [16]). An important question is how to combine different media (video, text, avatars) to reach a maximum of comprehensibility and comfort for people with different degrees of reading and signing skills. The ultimate question in the avatar domain is, however, how nonmanual components can be automatically integrated into existing systems and how that improves comprehensibility which will in turn affect overall acceptance.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] G. I. Balch and D. M. Mertens. Focus group design and group dynamics: Lessons from deaf and hard of hearing participants. *The American Journal of Evaluation*, 20(2):265 – 277, 1999.

[2] A. Braffort. Research on computer science and sign language: Ethical aspects. In *Gesture and Sign Language in Human-Computer Interaction*, LNAI 2298, pages 3–12. Springer, 2002.

[3] A. Cavender, R. E. Ladner, and E. A. Riskin. MobileASL: Intelligibility of sign language video as constrained by mobile phone technology. In *Proc. 8th intl. ACM SIGACCESS conf. on computers and accessibility (ASSETS)*, pages 71–78. ACM, 2006.

[4] A. C. Cavender, J. P. Bigham, and R. E. Ladner. ClassInFocus: Enabling improved visual attention strategies for deaf and hard of hearing students. In *Proc. 11th intl. ACM SIGACCESS conf. on computers and accessibility (ASSETS)*, pages 67–74. ACM, 2009.

[5] F. W.-l. Ho-Ching, J. Mankoff, and J. A. Landay. Can you see what i hear? the design and evaluation of a peripheral sound display for the deaf. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI)*, pages 161–168. ACM, 2003.

[6] J. A. Holt. Demographic, Stanford achievement test - 8th edition for deaf and hard of hearing students: Reading comprehension subgroup results. *Amer. Annals Deaf*, 138:172–175, 1993.

[7] M. Huenerfauth. A linguistically motivated model for speed and pausing in animations of american sign language. *ACM Trans. Access. Comp.*, 2(2):1–31, 2009.

[8] M. Huenerfauth and V. L. Hanson. Sign language in the interface: Access for deaf signers. In C. Stephanidis, editor, *The Universal Access Handbook*, chapter 38, pages 1–18. CRC Press, 2009.

[9] M. Huenerfauth, L. Zhao, E. Gu, and J. Allbeck. Evaluating american sign language generation by native ASL signers. *ACM Trans. Access. Comp.*, 1(1):1–27, 2008.

[10] J. R. Kennaway, J. R. W. Glauert, and I. Zwitserlood. Providing signed content on the internet by synthesized animation. *ACM Trans. Computer-Human Interaction*, 14(3):15–29, 2007.

[11] M. Kipp, A. Heloir, and Q. Nguyen. Sign language avatars: Animation and comprehensibility. In *Proc. 11th Intl. Conf. on Intelligent Virtual Agents (IVA)*. Springer, 2011.

[12] V. Lombardo, F. Nunnari, and R. Damiano. A virtual interpreter for the Italian sign language. In *Proc. 10th Intl. Conf. on Intelligent Virtual Agents (IVA)*, pages 201–207. Springer, 2010.

[13] T. Matthews, J. Fong, and J. Mankoff. Visualizing non-speech sounds for the deaf. In *Proc. 7th intl. ACM SIGACCESS conf. on computers and accessibility (ASSETS)*, pages 52–59. ACM, 2005.

[14] R. Nishio, S.-E. Hong, S. König, R. Konrad, G. Langer, T. Hanke, and C. Rathmann. Elicitation methods in the DGS (German Sign Language) corpus project. In *Proc. of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, 2010.

[15] C. Padden and T. Humphries. *Deaf in America: Voices from a Culture*. Harvard University Press, 1988.

[16] M. Sheard, S. Schoot, I. Zwitserlood, M. Verlinden, and I. Weber. Evaluation reports 1 and 2 of the EU project essential sign language information on government networks, Deliverable D6.2, March 2004.

[17] K. Shinohara and J. Tenenberg. Observing Sara: A case study of a blind person's interactions with technology. In *Proc. 9th intl. ACM SIGACCESS conf. on computers and accessibility (ASSETS)*, pages 171–178. ACM, 2007.

[18] J. J. Tran, T. W. Johnson, J. Kim, R. Rodriguez, S. Yin, E. A. Riskin, R. E. Ladner, and J. O. Wobbrock. A web-based user survey for evaluating power saving strategies for deaf users of mobileASL. In *Proc. 12th intl. ACM SIGACCESS conf. on computers and accessibility (ASSETS)*, pages 115–122. ACM, 2010.