Conference & Workshop on Assistive Technologies for People with Vision & Hearing Impairments
Assistive Technology for All Ages
CVHI 2007, M.A. Hersh (ed.)

END-USER INVOLVEMENT IN ASSISTIVE TECHNOLOGY DESIGN FOR THE DEAF – ARE ARTIFICIAL FORMS OF SIGN LANGUAGE MEETING THE NEEDS OF THE TARGET AUDIENCE?

S. Naqvi

Department of Computing, Goldsmiths College,
Goldsmiths Collge, University of London, New Cross, London, SE14 6NW, United Kingdom (UK)
Tel: +44(0)20 7919 7853 Fax: +44(0)20 7919 7462
Email: s.naqvi@gold.ac.uk

Abstract: A challenge facing individuals/organisations working with the Deaf community, is how to ensure that information is accessible to a community that has specialist language needs. Technologists have developed artificial digital representations of sign language that aim to communicate information to the Deaf community, such as video, animation and sign language notation systems. However these systems have received very mixed reviews and some are often rejected by the Deaf community. So what is it that makes one system more appropriate than another? In this paper we aim to use a novel approach to develop our understanding of how these systems are perceived, by comparing different well established digital sign language systems. This can highlight strengths and weaknesses for such systems to have a higher chance of success in the Deaf community. The paper will present the background and motivation, the experimental design, results and conclusions.

Keywords: Accessibility, Artificial Sign Language Systems, Digital Sign Language Systems, Assistive Technology

1. Introduction

With an increase in demand to facilitate the Deaf (big D and little d deaf are different communites, big D is culturally Deaf and very proud, whereas lower case d deaf are not heavily associated with the Deaf community and culture, for this research we are focusing on big D Deaf community) and hard of hearing in the often preferred language of British Sign Language (BSL). Several different systems have been made in order to facilitate this need. This has bought about the development of several different artificial forms of sign language which are based on sign languages themselves. The aim of such systems is to be closer to the natural language of the Deaf, use an efficient and transferable format and ease of reproduction. These systems will be referred to as Digital Representations of Sign Language (DRSL). In this research conducted these systems have been categorised as follows; notation (symbolic depictions of sign language (Sutton, V. 2007)), animation (artificial characters that can sign (Stories in the Air 2007)) and streaming video systems (filmed and edited sequences of sign language (Inclusive Learning Scotland 2007,)). Developing DRSL to assist in the distribution of deaf friendly information via the digital domain has created a variety of communication systems (Cox, S. et al 2002, J. Ohene-Djan, et al. 2003, J. Ohene-Djan, et al. 2004). Some systems are widely used while others receive very mixed reviews and are being compared to older systems thus there is no real effective transition to new and more advanced and effective technologies in terms of DRSL systems. The major criticisms arising around the systems lack of understanding of sign language and their inability to communicate affectively in a visual style (S. Naqvi, 2006).

Figure 1, Examples of (from left to right) Animation, Notation and Video representations of DRSLs

Researchers have postulated over what parts of these systems are unsuccessful and continue to redesign and generate new systems. However little to no research (to the authors knowledge) has been done in order to compare newer DRSLs to older ones, which could help identify the strengths and weaknesses in both and to move forward in development and design.

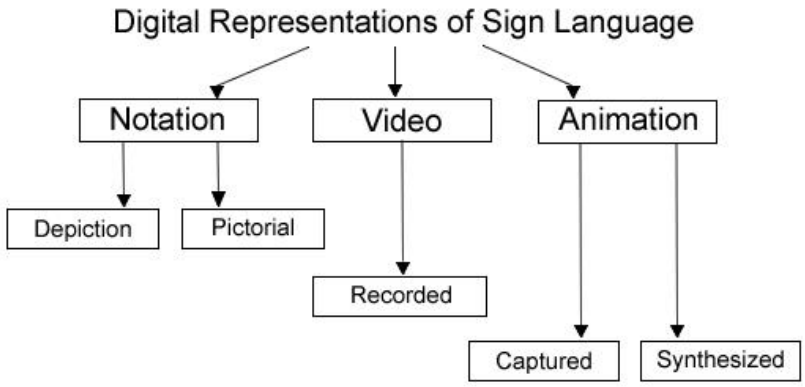

Figure 2, Breakdown of current categories of DRSLs

The research presented in this paper provides a novel approach to a comparative study of these systems against each other, in order to identify which linguistic components of BSL (the sign language used by the British Deaf community) are missing in these systems.

2. Background and Motivation

Traditionally video systems have been used to document and distribute information in sign language, this is a very expensive process in terms of time, expertise and facilities needed. We can record sequences that do not need to be filmed again such as manual instructions etc, however with more dynamic data such as weather reports, traffic updates and news flashes, how can we present the information in sign language if an interpreter is not available? One possible way is to put several different signs together in order to form new sequences of information. This can be done but proposes a new set of problems, different sequences concantenated together can loose the natural flow of signing and appear awkward. Also if the signer is dressed differently this can make the sequences visually confusing and discrepancies in different sequences will form.

In order to assist with Deaf accessible material notation systems are being used widely. They document signs and their movement, thus drawing a written version of sign vocabulary, and by putting these notations together one can write out sequences. This eliminates the problem of dissimilar video sequences and how they cannot be concatenated together, also providing a very flexible and transferable format, which will be technically viable to recreate and you can have the same notation style of writing in several systems, whereas with video you would need to have the same signer in the same clothes with the same look all the time. However the question arises can sign languages be drawn? By drawing sign languages the natural three-dimensional flow and movement, has been made flat in a written depiction of the language. This alphabet needs to be learnt in order to decode the drawing. Another possible solution for a dynamic system is Animation, this allows for the manipulation of an avatar (artificial human) to be puppeteered into signing whatever is programmed in.

The clothing, look, age, appearance, and styling are more superficial layers which can be manipulated. This will allow for several different signs to be programmed and joint together with greater technical ease and also providing the basis for more dynamic creation of sign languages. However do these artificial people compare well enough to real people signing?

These systems are different in their approaches to communication, however they are used and receive a variety of different opinions from the Deaf community. So the following questions are posed? For the static and dynamic creation of artificial sign language sequences which systems are more acceptable by the Deaf community, and what characteristics must be present in order to ensure the success of a system. The following research has attempted to understand how the Deaf community perceives these three systems, in the context of static and real-time information delivery. Each system demonstrates it's skill in all these contexts and were evaluated against each other. The static and dynamic sequences were evaluated against each other in terms of their linguistic ability.

3. Experimental Design

Four pilot studies were completed before this final approach was taken, in this paper we will only show the final design.

3.1 Hypothesis

The type of digital representation of sign language (i.e. avatar, video and notation systems) used in different information contexts (i.e. static, real-time) will determine higher acceptance rates of the systems and ultimately the efficiency and effectiveness of the information delivery. We aim to test the hypothesis that there is a difference in the perception of a DRSL in different information delivery contexts.

3.2 Methodology

The participants were given a form to complete, which gathered information about their levels of signing proficiency. Participants were then shown presentations that were samples of real-time and static productions of DRSL's. The DRSL's showed two different sets of information in order to avoid learning affects.

3.3 Materials

The user was presented with two information categories of presentations, which were, "static": which is information that is not changed often, such as a user manual or regular train times, and "real-time": this type of information is subject to change, such as venue changes for meetings to the reporting of changing weather conditions. In these two information categories the following sequences were shown

- Static:
    - Avatar: Pre-recorded sequence, which played an avatar clip
    - Video: Pre-recorded sequence, which played a streaming video clip
    - Notation: Pre-recorded sequence, which displayed pre-arranged notation graphics
- Real-time
    - Avatar: which played from the software
    - Video: which was a series of video clips concantenated together
    - Notation: a set of graphics that were presented on demand

3.4 Procedure

The procedure taken in the experimental conditions were:

- Introduction: The participants were called in one at a time and given general information, including clarifiation that the experiment was not a test, respondents should not worry about organisations involved, and that this is an independent study using dfferent technologies. They were also informed that they may stop the experiment at any time and they can remain anonymous. Before the participants could proceed they were asked their age, gender and contact information.
- Familiarization: The experiment was explained to the participants before the sample sequences were shown.
- Presentation: Participants were shown three presentations that were part of the two information categories (static and real-time). Each information category showed the three DRSLs (video, avatar, and notation). Each DRSL sequence was shown as many times as needed, until the participant was happy and ready to evaluate the presentation. The presentation was shown in different orders, to avoid primacy effects.
- Information gathering: Each participant was asked questions regarding the presentations they saw. They were asked if they understood what was said, and how they rated the presentation. Participants were also asked to fill in another questionnaire which examined personality types, and if this had an affect on the perception of such systems.
- Questions and answers were at the end of the session

3.5 Data Collection Method Used

The data was collected through questionnaires, and video filming which was later analysed.

4. Results

Static and real-time presentations were tested against each other and questions were asked about how the presentations were perceived. The categories of assessment in terms of BSL linguistics were: hand shape, morphology, distance of the arm from the body, lip movement, facial expression, correct sentence BSL structure, correct placement and correct signing context.

4.1 Data collection method used

Overall Results of Systems	N	Minimum	Maximum	Mean	Std. Deviation
Animation Linguistics	20	1.75	5.00	3.5563	.91180
Notation Linguistics	19	2.25	5.00	4.1711	.93096
Video Linguistics	20	1.00	3.00	2.0313	.69995
Valid N (listwise)	20				

NB: the scale was 1 = Excellent to 5 = Poor.
NB: a participant refused to rate the notation system in terms of linguistics as they strongly felt it was not in keep with sign language.

Figure 3 Descriptive Statistics For Static Presentations of DRSL's

Overall Results of Systems	N	Minimum	Maximum	Mean	Std. Deviation
Animation Linguistics	20	1.50	5.00	3.3875	1.01137
Notation Linguistics	20	1.75	5.00	4.2625	.88286
Video Linguistics	20	1.38	4.38	2.3250	.78995
Valid N (listwise)	20				

NB: the scale was 1 = Excellent to 5 = Poor.

Figure 4 Descriptive Statistics For Real-time Presentations of DRSL's

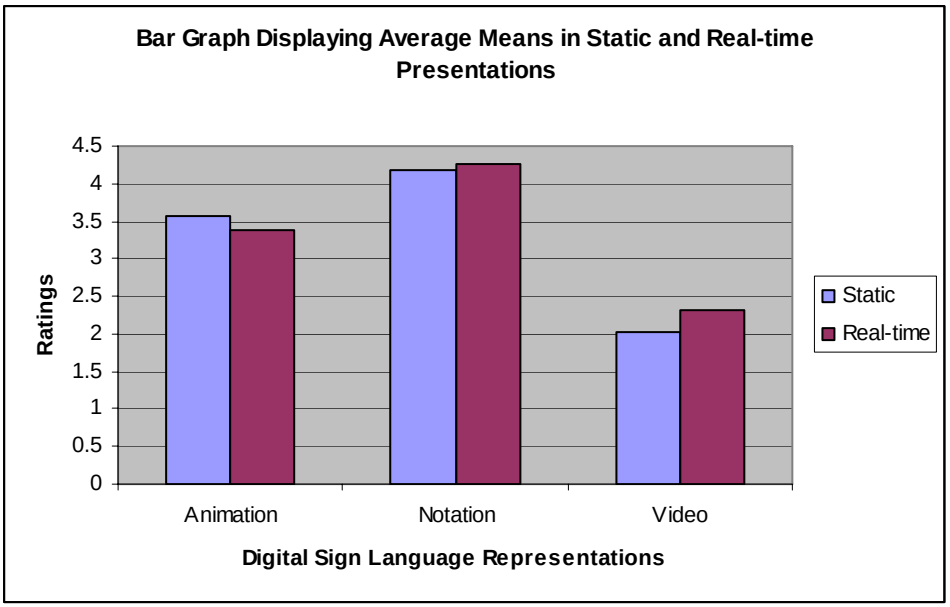

Figure 5 Bar Graph Displaying Average Means in Static and Real-time Presentations

4.2 Inferential statistics

Further inferential analysis was conducted by a series of T-tests. We noted that the only result that showed statistical significance was in the animation digital representation of sign language, under the linguistic category of hand shape. It was noted that the overall means were statistically significant; the mean for the static presentation of animation was 3.2 and the mean for the real-time presentation was 3.8, $t(19) = 2.11$, $p<0.05$. This was on a scale where 1 was rated excellent and 5 was rated as poor. It can be suggested by the means that the static presentation was rated better than the real-time presentation.

The results have been listed in the table below, as you can see from the average means the systems were rated quite poorly on the scale, where 1 was excellent and 5 was poor:

Presentation Mode	**Category**	***r***	**Significance**	**Average Mean**
Animation	Morphology	.611	0.004	(3.4+3.6)/2 = 3.5
Animation	Lip Movement	.583	0.007	(4.3+4.1)/2 =4.25
Animation	Facial Expression	.384	0.095	(4.2+4.1)/2 =4.2
Animation	Correct sentence BSL structure	.436	0.054	(3+3.2)/2 =3.05
Animation	Linguistics	.601	0.005	(3.4+3.6)/2 =3.5
Notation	Handshape	.606	0.006	(4.3+4.1)/2 =4.2
Notation	Morphology	.732	0.000	(4.1+4)/2 =4.05
Notation	Distance of the arm from the body	.430	0.066	(4.3+4.3)/2 =4.3
Notation	Lip movement	.692	0.004	(4.4+4.4)/2 =4.4
Notation	Facial Expression	.704	0.001	(4.2+4.5)/2 =4.35
Notation	Correct sentence BSL structure	.466	0.002	(4.3+4)/2 =4.15
Notation	Linguistics	.699	0.001	(4.3+4.2)/2 =4.25

NB: correlations are between static and real-time presentations of DRSLs for the respective variables shown in the table. Means are shown in order to indicate that the high correlations and are not withstanding the variables which were generally high (high ratings indicates poor ratings)

Figure 6 Table of results showing correlations between static and real-time presentations of DRSLs

It was observed that although the digital representations of sign language were different in static or real-time mode, they had significantly high correlations. In the digital representation of animation, when static and real-time presentations were shown, the following categories had significant correlations, likeability r = .75, usability r = .67, linguistics r = .60, under linguistics the following were observed; morphology r= .61, lip movement r = .59, facial expression r = .39, correct sentence BSL structure r = .44. In the digital representation of Notation the following correlations were noted, likeability r = .75, acceptability r = .72, comprehension r = .817, linguistics r = .70, under linguistics the following were observed hand shape r = .61, morphology r = .73, distance of the body from the arm r = .43, lip movement r = .70, facial expression r = .70, correct sentence BSL structure r = .47.

There were several high correlations between the variables used in the paired sample t-tests between static and real-time presentations of each digital representation of sign language. The main significant differences and correlations were found only in animation and notation, video did not show any statistically significance in terms of difference or correlation.

5. Conclusions

From the results we can see that video is still the most popular DRSL, and there are particular characteristics of BSL linguistics that appear to be missing from animation and notation systems. The correlations between static and real-time presentations in both animation and notation systems were: morphology, lip movement, facial expression, correxct sentence BSL structure, in particularly with notation systems was the distance of the arm from the body. It was also noted that throughout the experiment participants remarked on the systems lack of “emotion” and that is why they could not understand what was being communicated in Animation and Notation systems. It was also notable that in one particular notation image there was a face drawn into the notated sign, and participants often remarked when seeing this image that it “made sense”. But with the other notations without faces drawn in participants simply said “this is not clear” and “I cannot understand what this word means”. It was also interesting to note that although the video sequence for real-time generation of BSL was a series of concantenated signs (where the signer raised his hands, signed a word and droped his hands back down) had better ratings than the avatar that produced smooth flowing signing without dropping hands. So although the animation provided a more fluid form of signing and a more “technically appropriate” solution for real-time sign language generation, the smooth flow of handshapes and body movement was not enough and the biggest thing that participants relied upon for communication was the face. This suggests that the hypothesis tested was not true, and that particular components of BSL need to be present in order to improve the systems acceptance and use by the Deaf community.

References

Cox, S., et al (2002), Tessa, a system to aid communication with deaf people, *In Proc. Assets ‘02, ACM Press*, New York - USA, 2002.

Inclusive Learning Scotland (2007), Science Signs, http://www.ltscotland.org.uk/inclusiveeducation/findresources/bslsciencesigns/BSLinteractive/bslscience.asp

Ohene-Djan, J. et al. (2003), A Personalisable Electronic Book for Video-based Sign Language Education, *In Proc IFETS, IEEE, pp. 86-99.*

Ohene-Djan, J. et al. (2004), Mak-Messenger and Finger-Chat, communications technologies to assist in teaching of signed languages to the deaf and hearing, *In Proc. ICALT, IEEE*, Nigata-Japan, pp. 744-746.

Stories in the Air (2007). *British Sign Language Dictionary Online*, http://www.learnbsl.org/

S. Naqvi, (2006). Experimental research design for the Deaf: approaches for obtaining accurate end-user feedback, *In Proc.CVHI,* pp. CD– ROM, euro-Assist-CVHI-4.

Sutton. V. (2007). SignWriting, http://www.signwriting.org/

Acknowledgements: I would like to thank, Dr. James Ohen-Djan, Dr. Herbert Blumberg, Valerie Sutton, Royal National Institue for the Deaf in particular Dr. Mary Sheard and Antony Rabin. Edward J Richards, Penny Beschizza and all the students of Southwark College, Herbert Klein, Shahab Reza and Deafax in the support and evaluation of this work.