

2013

Manual Evaluation of synthesised Sign Language Avatars

Robert G. Smith Mr

Technological University Dublin, robert.smith@tudublin.ie

Brian Nolan Dr.

Technological University Dublin

Follow this and additional works at: <https://arrow.tudublin.ie/itbinfocon>

Digital Part of the [Communication Technology and New Media Commons](#), [Disability Studies Commons](#), and [the Other Communication Commons](#)
Network

Logo

Recommended Citation

Smith, R.G. & Nolan, B. (2013). Manual Evaluation of synthesised Sign Language Avatars. *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*, pg. 57, ACM.

This Conference Paper is brought to you for free and open access by the Computational Functional Linguistics at ARROW@TU Dublin. It has been accepted for inclusion in Conference Papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact arrow.admin@tudublin.ie, aisling.coyne@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 4.0 License](#)

Funder: ITB

Footer logo

Manual Evaluation of synthesised Sign Language Avatars

Robert Smith, Brian Nolan
Department of Informatics,
Institute of Technology, Blanchardstown,
Dublin, Ireland.
{robert.smith, brian.nolan}@itb.ie

ABSTRACT

The evaluation discussed in this paper explores the role that underlying facial expressions might have regarding understandability in sign language avatars. Focusing specifically on Irish Sign Language (ISL), we examine the Deaf community's appetite for sign language avatars. The work presented explores the following hypothesis: Augmenting an existing avatar with various combinations of the 7 widely accepted universal emotions identified by Ekman [1] to achieve underlying facial expressions, will make that avatar more human-like and consequently improve usability and understandability for the ISL user. Using human evaluation methods [2] we compare an augmented set of avatar utterances against a baseline set, focusing on two key areas: comprehension and naturalness of facial configuration. We outline our approach to the evaluation including our choice of ISL participants, interview environment and evaluation methodology.

General Terms

Design, Human Factors, Languages.

Keywords

User-centered design, Deaf, Sign Language synthesis, Emotion, Avatar, SiGML, HamNoSys, Accessibility, Disability, HCI.

1. INTRODUCTION

ISL is the indigenous language of the Deaf Community in Ireland, standing apart from English and Irish. Unlike spoken language, signed languages have multiple articulators designated: Manual features (MF) which are the hands/arms and non-manual features (NMF), everything else. Emotion and prosody are expressed in SL primarily through NMFs [3], which are widely accepted to carry up to 70% of a sign's meaning¹ and this, therefore, makes emotion a significant factor in the credibility and acceptance of an avatar. The average reading age of Deaf school leavers is comparable to that of an 8-9 year old hearing child [5]. Thus there is a requirement for communication materials in a sign language format yet the costly production of sign language video means

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

Copyright is held by the owner/author(s).

ASSETS '13, Oct 21-23 2013, Bellevue, WA, USA
ACM 978-1-4503-2405-2/13/10.

that these materials are limited. Synthesised sign language avatars are a cost effective solution to this requirement. This paper outlines how an existing synthesised avatar framework is currently being used to evaluate comprehension levels of signing avatars amongst a portion of the Irish Deaf community. Particularly, the study investigates the effect of adding emotional facial expressions (EFEs) and the advantages, if any, of a human-like avatar versus a caricature-like avatar.

2. EVALUATION FRAMEWORK

In order to conduct any form of evaluation one must first have an avatar in place along with some synthesised output. With this in mind a subset from the well-established Signs of Ireland (SOI) corpus [6] and the JASigning [7] synthesised sign language avatar system were chosen as the instruments to work with. These technologies are discussed further in the subsequent sections.

3. BASELINE SYSTEM

During the EU funded ViSiCAST and eSIGN projects the JASigning framework was developed [7]. This very well established framework, the driving force behind avatars such as Anna (figure 1), was developed with a modular structure such that researchers outside of the initial projects can easily pick up the technology and build on it. With the JASigning framework: a transcriber may represent a signed utterance at the phonetic level using the Hamburg notation system, HamNoSys [8]. This notation is then represented by the framework as a markup language called SiGML (Signing Gesture Mark-up Language) [9]. SiGML defines a set of XML tags for each iconic symbol in HamNoSys. Using this information the animation may then be synthesised, and rendered in real-time.

3.1 Improving the baseline system

Taking advantage of the frameworks modular structure it was possible to alter a number of external XML and property files without having to delve into a labyrinth of code. This allowed us to target the aspects of the avatars, which effect the facial movement. The ARToolkit [10] was used to create new facial morphs that represent Ekman's 7 universal emotions: happiness, sadness, anger, disgust, contempt, fear and surprise. At this point it was necessary to manually add the EFEs markup to each SiGML file for each avatar.

4. EVALUATION

A manual evaluation was undertaken with 15 sign language users over a two-day period on site at the newly developed Deaf village of Ireland (DVI). The evaluation was designed such that all participants were native ISL users and a demographic balance was

¹ Anecdotal evidence from conversations with ISL tutors and our ISL animation evaluators

achieved. Barriers such as different levels of technical knowledge and pre-formed opinion of the technology were identified. Some barriers, like communication, for example, were overcome with the support of a certified ISL interpreter. A set of 5 digital videos varying in duration from 9 seconds to 73 seconds were viewed by the participants. Each saw the videos in a different sequence, which was derived using a Latin square model to avoid learning. No participant saw all 5 videos. The longest video was always shown in isolation or with one other to prevent fatigue in the participant. Each of the 5 videos comprised of one of four avatars: (a.) Anna, a 'human looking' avatar with baseline encoding, (b.) Luna, a caricature avatar with baseline encoding and both (c.) Anna and (d.) Luna enriched with EFEs. Making a total of 20 avatar videos. After watching each video the participants were asked a number of comprehension questions as well as being asked to score their own comprehension of the video content on a scale of 0-5. During a trial run of the evaluation it became obvious that some context was required and each video would need to be watched a second time, therefore, the same set of questions were asked after each viewing in a bid to track the level of comprehension after each pass.

5. RESULTS

Indications are that participants, when directly asked, underestimated their own comprehension on each avatar video shown. On average participants self-scoring for each avatar is considerably lower than the score achieved in the comprehension exercise. In the case of avatars that had been enriched with EFEs, the self-applied score was 14% lower. At the other extreme, in the case of the avatar Anna (with and without EFEs), the score was 44% lower. This indicates that the participant's perceived comprehension is substantially lower than their actual comprehension, which may be one reason for the low uptake of this technology amongst the Deaf community. Surprisingly, participants understood 62% of the content delivered through the baseline avatars yet when EFE was added the comprehension level fell to 60%. This would seem to indicate that instead of improving comprehension, the addition of EFE had a negative effect albeit marginal. A further breakdown of the results gives a clearer picture as to how each of the four avatars performed. Anna with EFE recorded a higher comprehension score than Luna with EFE scoring 64% and 54% respectively. Anna also scored higher than Luna with the baseline encoding, with score of 63% and 59% respectively. Again we note the gulf between the EFE and baseline avatars. These results demonstrate that the addition of EFEs for comprehension was more successful with the 'human looking' avatar than with the caricature avatar. In addition, these results also confirm that regardless of EFEs, Anna was the easier avatar to comprehend.



Figure 1 Avatars Luna and Anna

6. CONCLUSION

The primary focus of the evaluation was to ascertain whether or not the addition of emotional facial configuration increased the understandability of a signed utterance. The results presented here would indicate that this is not the case. In fact, the addition of

EFEs made very little impact with the score for the baseline avatars and the EFEs augmented avatar almost identical, overall having a marginally negative effect of -2%. Also evident from the results is the higher comprehension levels achieved with the avatar Anna. Anna was designed to be as close to human looking as possible while using lower levels of 3D data for speedy rendering. This result could have a significant impact on future development of sign language avatars and their facial configuration. Commonly, participants commented that Anna looked quite the serious avatar and that Luna's large eyes and long fingers may be better suited for children. Regardless of preference, however, results show that participants understood Anna better than Luna. Participants highlighted a number of attributes of linguistic importance that performed badly including emotion and NMF and natural movement. One must surmise that these linguistics attributes are directly linked to the participants' comprehension and indeed the perceived comprehension scores. Although the average comprehension scores indicate only a minor effect of EFEs, these attributes are desired by the Deaf community and furthermore, if implemented more comprehensively, are required to improve comprehension past 60%.

7. ACKNOWLEDGMENTS

This research is supported by: The School of Informatics and Engineering at the Institute of Technology Blanchardstown. We would like to acknowledge the support of our colleagues at UEA and the efforts of the Irish Deaf Society, particularly Susan Whelan and the participants who freely offered their time for this evaluation.

8. REFERENCES

- Ekman, Paul. Basic Emotion. In *Handbook of cognition and emotion*. Wiley, UK, 1999.
- Huenerfauth, Matt, Zhao, Liming, Gu, Erdan, and Allbeck, Jan. Evaluation of American Sign Language Generation by Native ASL Signers. *Transactions on Access Computing*, Vol. 1, 1 (2008), Article 3.
- Matthews, P. *The Irish Deaf Community, Survey Report, History of Education, Language and Culture*. The Linguistics Institute of Ireland, Dublin, 1996.
- Leeson, L. *Aspects of Verb Valency in Irish Sign Language*. Trinity College, Dublin, 2001.
- Conrad, R. *The Deaf School Child: Language and Cognitive Function*. HarperCollins Publishers, London, 1979.
- Leeson, L and Nolan, B. Digital Deployment of the Signs of Ireland Corpus in Elearning. In *3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora* (Marrakech 2008), 112-121.
- Kennaway, R. Experience with and requirements for a gesture description language for synthetic animation. In A. Camurri and G. Volpe (eds) *Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop* (Genova 2003), 300-311.
- Prillwitz, S. *HamNoSys Version 2.0: Hamburg notation system for sign languages: An introductory guide*. Signum, 1989.
- Elliott, R, Glauert, J, Jennings, V, and Kennaway, R. An overview of the SiGML notation and SiGMLSigning software system. (2004), 98-104.
- Vince Jennings, Ralph Elliott, Richard Kennaway, John Glauert. Requirements For A Signing Avatar. In *LREC, 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies* (2010).