Models of Linguistic Facial Expressions For American Sign Language Animation

Hernisa Kacorri
The City University of New York (CUNY)
The Graduate Center
Doctoral Program in Computer Science
365 Fifth Ave, New York, NY 10016
hkacorri@gc.cuny.edu

Abstract

Facial expressions are a necessary and important part of the grammar of American Sign Language (ASL) sentences. However, current computational linguistic software cannot produce a huge portion of ASL constructions that include facial expressions. Here we summarize our research motivations, proposed plan, and potential contributions in the accessibility field. Our research aim is to formulate signer-independent models of facial movements for various types of linguistically meaningful facial expressions in ASL animation. This project summary also includes a brief description of current progress and future research plans.

Introduction

Software to generate American Sign Language (ASL) animations yields significant accessibility benefits for many signers with lower levels of written language literacy in the USA [Huenerfauth, M. and Hanson, V., 2009]. However, it is still challenging for modern ASL animation software to support accurate and understandable signing virtual human characters.

State of art ASL animation tools focus mostly on the accuracy of manual signs, not facial expressions. However, facial expressions such as furrowed or raised eyebrows, pursed lips, and movements of the head and upper body reveal linguistically significant information in ASL. When applied they can also indicate the grammatical status of phrases or entire sentences. E.g., the only way that a yes/no-question, a wh-question (question phrases like 'who,' 'what,' etc.), a negation or a topic is conveyed is with nonmanual components such as the face and head movement. Figure 1 illustrates an example where two ASL sentences with the same sequence of signs performed by hands are interpreted differently based on the accompanying facial expression.

Thus, the production of grammatical facial expressions and head movements in coordination with specific manual signs is crucial for the interpretation of ASL sentences. There is a significant difference in deaf users' comprehension of ASL animations when linguistically and emotionally meaningful facial expressions are supported [Huenerfauth et. al., 2011].

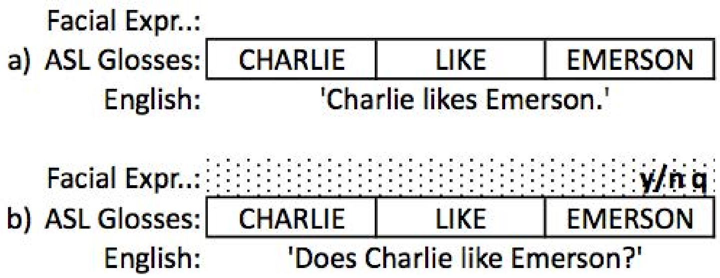

Figure 2: Differentiating an ASL statement (a) from an ASL question (b) based on the associated facial expression.

Suppose a detailed input is provided to an ASL animation system. This would include: (i) a script specifying the sequence of the ASL glosses; (ii) the associated facial expressions; and (iii) the corresponding facial expressions' start- and end-glosses (as shown in Fig.1). There are still a number of challenges to overcome, each of which leads to some research questions (Q1-Q4 below).

Performing the facial expression: A facial expression is not just a mask but also a dynamic movement requiring careful "choreographing" of numerous parameters of facial landmarks. *[Q1]: How should the face be articulated to perform, with accuracy, the linguistically meaningful facial expressions that are part of ASL grammar?*

Synchronizing the facial expression to hand-movements: Manual signs are precisely coordinated with associated nonmanual expressions (e.g., in yes/no-questions, eyebrows typically rise in advance of initial co-occurring manual signs; they return to neutral position slightly in advance of the end of the final manual sign.) Failure to do so can result in misinterpretations or slower comprehension. Synchronization of facial expression intensity with manual signs is also complex [Neidle et. Al., 2000]. *[Q2]: How should the facial expressions and hand movements be temporally coordinated?*

Transitioning between sequential facial expressions: This usually requires complex transition rules to describe the blending between facial expressions, not captured by simple interpolation (e.g., in Fig. 2 yes/no question marking follows topic marking). *[Q3]: How should the onsets, offsets, and transitions of these movements be produced?*

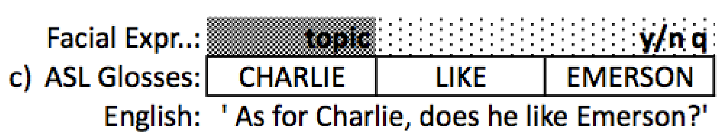

Figure 3: Sequential facial expressions in an ASL sentence.

Performing multiple simultaneous facial expressions: An example is shown in Fig. 3 where a yes/no-question marking co-occurring with a negation marking follows a topic marking. Learning the underlying rules for overlaying one facial expression onto another is crucial. *[Q4] How should multiple simultaneous facial expressions be combined?*

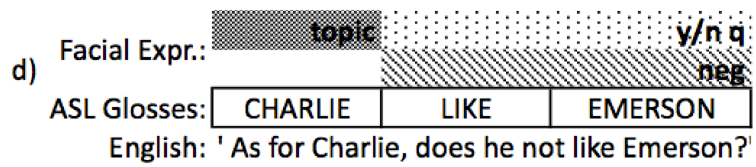

Figure 4: Simultaneous facial expressions in an ASL sentence.

No ASL animation tool fully supports the above. Therefore, these features could extend the state of the art in ASL animation technology and provide measurable benefits for deaf users.

Research Goals

This project goal is to raise the understandability and perceived quality of the ASL animations through the inclusion of facial expressions; specifically, we will design computational models for use in ASL-animation software and ask native signers to evaluate the result. Wh-question, yes-no question, rhetorical question, negation, and topicalization are the facial expressions to be studied. Our work will be based on a set of linguistically annotated video data collected from human signers and provided by collaborators of at Boston University (BU). Other collaborators at Rutgers University will next use computer vision techniques to track the facial landmarks of the humans in the videos. Our animation platform is based on an open source software library/toolkit, EMBR [DFKI, 2012], produced by collaborators at DFKI. EMBR allows for detailed control of an animated human character, and we are helping to extend it with standard MPEG-4 Face Parameters [SNHC, 1998]. This project tentative research plan is summarized as follows:

- Creating parameterized facial models based on MPEG-4.
- Training machine-learning models based on linguistic hypotheses from BU team for use in animation generation. Features that are critical to the timing and intensity of facial expressions provided by the Rutgers team will be considered.
- Building ASL scripting infrastructure, prepopulating the infrastructure with animated ASL signs, and creating stimuli with the assistance of native ASL signers.
- Designing and conducting experiments with native ASL signers to evaluate animations based on our models.
- Revising and retesting our hypothesized models iteratively, based on the results of experiments with native signers.

Current Progress

Initial research conducted during the past year includes: (i) investigation of the newly available EMBR toolkit for synthesizing the animations in this research project; (ii) software development, used this summer by interns at our lab to create new ASL signs; (iii) conduct of studies on the effects Effect of presenting videos of actual human signers as a baseline during American Sign Language animation evaluation studies with native ASL participants [Lu and Kacorri, 2012]; (iv) familiarization with ASL and basic aspects of American Deaf Culture.

During the summer, the lab provides summer research opportunities for deaf students [Huenerfauth, 2010]. This summer, we worked with one undergraduate student and three high school students who are all native ASL signers. Together, we (a) populated the ASL animation lexicon by building new signs, (b) designed sentences/stimuli for use in upcoming experiments, and (c) created animation for these sentences in EMBR as experimental stimuli for the project.

Conclusions

Limitations in the grammatical-correctness and naturalness of facial expressions have held back the understandability of ASL animations. The main contribution of this project is the creation of high quality models of the movement of virtual humans in ASL animation. The proposed techniques should also be applicable to animations of other sign languages. Other contributions of this work include: (1) providing an evaluating methodology for research studies on the understandability and naturalness of ASL animations with facial expressions, (2) contribute with a collection of empirical data on facial expression preferences and comprehension rates of ASL animations by native ASL signers, and (3) give motivation for future computational linguistic work on ASL.

This material is based upon work supported in part by the US National Science Foundation under award number 1065009.

References

DFKI. 2012. Homepage. http://embots.dfki.de/EMBR/

Huenerfauth, M. 2010. Participation of High School and Undergraduate Students who are Deaf in Research on American Sign Language Animation. ACM SIGACCESS Accessibility and Computing newsletter. New York: ACM Press. Issue 97 (June 2010).

Huenerfauth, M., Hanson, V. 2009. Sign language in the interface: access for deaf signers. In C. Stephanidis (ed.), Universal Access Handbook. NJ: Erlbaum. 38.1-38.18.

Huenerfauth, M., Lu, P., and Rosenberg, A. 2011. Evaluating Importance of Facial Expression in American Sign Language and Pidgin Signed English Animations. In Proceedings of The 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2011), Dundee, Scotland. New York: ACM Press.

Lu, P. and Kacorri, H. (2012, in press). Effect of Presenting Video as a Baseline During an American Sign Language Animation User Study. In Proceedings of The 14th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2012), Boulder, Colorado. New York: ACM Press.

Neidle, C., Kegl, D., MacLaughlin, D., Bahan, B., Lee, R.G. 2000. The syntax of ASL: functional categories and hierarchical structure. Cambridge: MIT Press.

SNHC. 1998. Information Technology – Generic Coding Of Audio – Visual Objects Part 2: Visual, In ISO/IEC 14496-2, Final Draft of International Standard. In ISO/IEC JTC1/SC29/WG11 N2502a, Atlantic City.

About the Author:

Hernisa Kacorri is currently enrolled in the Computer Science PhD program at The Graduate Center, City University of New York (CUNY). She obtained her bachelor and master degrees in Computer Science and Advanced Information Systems, respectively, from the University of Athens, Greece. Her current research focuses on computer accessibility, assistive technology, and human computer interaction. She is a research assistant at the Linguistic and Assistive Technologies Laboratory at Queens College, CUNY under the supervision of the Associate Professor Matt Huenerfauth.