Release of Experimental Stimuli and Questions for Evaluating Facial Expressions in Animations of American Sign Language

Matt Huenerfauth
The City University of New York (CUNY)
Computer Science Department, Queens College
65-30 Kissena Blvd, Flushing, NY 11367 USA
E-mail: matt@cs.qc.cuny.edu

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

Abstract

We have developed a collection of stimuli (with accompanying comprehension questions and subjective-evaluation questions) that can be used to evaluate the perception and understanding of facial expressions in ASL animations or videos. The stimuli have been designed as part of our laboratory's on-going research on synthesizing ASL facial expressions such as Topic, Negation, Yes/No Questions, WH-questions, and RH-questions. This paper announces the release of this resource, describes the collection and its creation, and provides sufficient details to enable researchers determine if it would benefit their work. Using this collection of stimuli and questions, we are seeking to evaluate computational models of ASL animations with linguistically meaningful facial expressions, which have accessibility applications for deaf users.

Keywords: American Sign Language, facial expression, non-manual signals, stimuli.

1. Introduction

Synthesis of American Sign Language (ASL) animations can provide benefits for deaf and hard-of-hearing people with lower levels of written language literacy (Huenerfauth, 2004a). This is underscored by the literacy rates of deaf adults in the United States on standardized testing (Traxler, 2000) and the large number of ASL users (over 500,000) in the United States (Mitchell et al., 2006). In prior experimental studies, we determined that the use of emotional and linguistically meaningful facial expressions in ASL animations significantly increased viewers' comprehension and perceived quality of animations (Huenerfauth, Lu, and Rosenberg, 2011). To produce an animation with natural facial expressions, a skilled animator and ASL signer could carefully control the face of the avatar on a fine-grained timeline, but such an approach is time-consuming and depends too much on the skills of the animator. Thus, a more automated solution is needed to minimize the required input in order to produce an animation; this minimal input script would include only the sequence of glosses, the type of facial expression needed, and the starting and ending glosses in the sentence when it should occur.

Many prior sign language animation systems lack sophisticated models in support of non-manuals, which are necessary to automatically synthesize clear and understandable facial expressions. There has been recent work by several groups to improve the state-of-the-art of facial expressions and non-manual signals for sign language animation, e.g.: Wolfe et al. (2011) and Schnepp et al. (2012) used linguistic findings to drive eyebrow movement in animations of interrogative (WH-word) questions with or without co-occurrence of affect. Schmidt et al. (2013) used clustering techniques to obtain lexical facial expressions. Gibet et al. (2011) used machine-learning methods to map facial motion-capture data to animation blend-shapes.

This paper presents a collection of stimuli to evaluate the perception and understanding of facial expressions in ASL animations. Section 2 describes the research project for which the stimuli were developed; section 3 provides basic information about the stimuli and briefly explains the linguistics of the facial expressions within each. Section 4 gives additional detail about how the stimuli and questions were engineered to measure the perception and comprehension of facial expressions. Section 5 describes how facial movements in the stimuli videos were identified and recorded, and section 6 describes prior studies that used some of these stimuli. Section 7 contains information about how to obtain the collection.

2. Our Research on ASL Animation

The goal of our ongoing research is to improve technologies for generating ASL animations through the inclusion of linguistically meaningful ASL facial expressions. We seek to develop computational models to generate facial expressions that convey grammatical syntax information such as topic, negation, rhetorical questions, WH-word questions, and yes/no questions (Kacorri, 2013). It is necessary to model how elements of the face move during ASL facial expressions, how these movements are timed in relation to the manual signs, and how these facial movements co-occur or segue into one another. In pursuit of this goal, our lab has begun to analyze linguistically annotated ASL videos (Liu et al., 2013) and automatically tracked facial landmarks in these videos (Yu et al., 2013) so that we may create signer-independent models that can generate grammatically correct ASL animations with facial expressions.

To evaluate our animation models, native ASL signers typically view our animations and answer subjective Likert-scale and comprehension questions (Huenerfauth, 2004b; Huenerfauth et al., 2007; Huenerfauth, 2008). Inventing stimuli and comprehension questions that effectively measure whether participants understand the information conveyed specifically by the model-driven face can be challenging. Several facial expressions affect

the meaning of ASL sentences in subtle ways (Kacorri, Lu, and Huenerfauth, 2013b) and often signers may not consciously notice a facial expression during an ASL passage (Huenerfauth, Lu, and Rosenberg, 2011; Kacorri, Lu, and Huenerfauth, 2013b).

During our multi-year project, we have experimented with different forms of stimuli design strategies to elicit ASL passages and comprehension questions that can measure whether the viewer has understood linguistic facial expressions correctly (Kacorri, Lu, and Huenerfauth, 2013b). After three years of user studies on ASL facial expressions that convey grammatical syntax information (Huenerfauth, Lu, and Rosenberg, 2011; Kacorri, Lu, and Huenerfauth, 2013a; Kacorri, Lu, and Huenerfauth, 2013b; Kacorri, Harper, and Huenerfauth, 2013), we have designed a collection of scripted ASL multi-sentence single-signer passages and corresponding comprehension questions that probe whether human participants watching these stimuli have understood the information that should have been conveyed specifically by the facial expressions. We are now sharing with the research community the set of stimuli and questions we have developed in support of our research on non-manual linguistic phenomena.

3. Overview of the Collection

This paper is the first announcement of the release of this stimuli collection, which includes: 48 ASL passages performed by a native signer; 192 comprehension questions (4 questions for each passage, each question performed by 2 native signers, male and female); a set of Likert-scale subject questions about the grammatical correctness, ease of understanding, and naturalness of movement of the passages; and a set of Likert-scale questions asking whether participants noticed specific categories of facial expressions. The collection consists of video recordings of a native ASL signer, ASL transcriptions of each passage, English translation of the ASL passages and comprehension questions as plaintext files, and two sets of questionnaires with the Likert-scale questions. The English translations of the ASL stories includes both the indented meaning when the ASL facial expression is performed correctly and a second ambiguous meaning when the facial expression is not correctly perceived by the person viewing the story.

Each stimulus focuses on a particular facial expression in one of the following categories listed below. Each is illustrated in Figure 1 and informally described below; please consult ASL linguistics references for more detailed explanations, e.g., (Neidle et al., 2000).

- Yes/No Questions: The signer raises his eyebrows while tilting the head forward during a sentence to indicate that it should be interpreted as a question.
- WH-Questions: The signer furrows his eyebrows and tilts his head forward during a sentence that should be interpreted as information-seeking, typically with a "WH" word such as what, who, where, when, how, which, etc.
- RH-Questions: The signer raises his eyebrows and tilts his head backward and to the side to indicate a question that should be interpreted rhetorically.

Figure 1: Still images taken from videos included in the stimuli collection described in this paper, with each image illustrating a moment when a particular facial expressions is occurring: (a) YN-Question, (b) WH-Question, (c) RH-Question, (d) Topic, (e) Negation, and (f) Emotional Affect (an example of anger is shown in this image).

- Topic: The signer raises his eyebrows and tilts his head backward during a phrase at the beginning of a phrase that should be interpreted as a topic.
- Negation: The signer shakes his head left and right during the verb phrase which should be interpreted with a negated meaning, often with the sign NOT.
- Emotional affect: These facial expressions are not linguistically governed, but they include several typical affective facial expressions that can indicate sadness, anger, frustration, etc. during a sentence.

The value of this collection is that the stories and questions were carefully engineered so that the participant must perceive and understand the facial expression in order to answer the comprehension questions correctly. For each stimulus, if the manual portion of the performance were considered alone (without the facial expressions), then there would be an ambiguity or an alternative semantic interpretation possible for the stimulus. Our comprehension questions have been designed to detect when a participant has misunderstood the stimulus due to the facial expression not being successfully perceived or understood. Thus, these stimuli can be used to evaluate the quality of automatic animation-synthesis systems for generating animations of ASL with facial expressions.

Table 1 provides a listing of the number of stimuli in the collection of each type.

Table 1: Collection Overview.

Type of facial expression	Number of stimuli (Average number of glosses per stimulus)	Codenames of these stimuli in the collection
Emotional Affect	8 stimuli (6.88)	E1, E2, E3, E4, E5, E6, E7, E8
WH-word Questions	9 stimuli (13)	W1, W2, W3, W4, W5, W6, W7, W8, W9
Yes/No Question	7 stimuli (9.29)	Y1, Y2, Y3, Y4, Y5, Y6, Y7
Topic	7 stimuli (10)	T1, T2, T3, T4, T5, T6, T7
Rhetorical Question	11 stimuli (11.82)	R1, R2, R3, R4, R5, R6, R7, R8, R9, R10, R11
Negation	6 stimuli (16.5)	N1, N2, N3, N4, N5, N6

4. Design of Stimuli and Questions

Prior to the design of the stimuli, a native ASL signer was given 6 categories of facial expressions and was introduced to premise that the passage must be ambiguous in its meaning if the facial expression were not understood. The native ASL signer invented, performed, and transcribed the ASL passages, and the passages were discussed and edited in collaboration with a team of other native ASL signers at the laboratory. Next, the two ambiguous meanings were translated into English sentences. Consulting the ASL transcription and the two ambiguous English translations, a second native ASL signer performed the ASL passages for the video recordings in our collection. Finally, linguistic researchers at our laboratory engineered the comprehension questions for each story such that they would receive different answers, depending on the perception and understanding of the facial expression. The collection includes a sample HTML form where the 4 comprehension questions are embedded in video format and the answers are collected on a 7-point Likert scale from "definitely no" to "definitely yes."

While researchers can access the full collection of stimuli and questions, this section explains a specific example of each category of stimuli to illustrate how each stimulus can have alternative interpretations, if the facial expression were not correctly understood.

4.1 Example: Topic

The following sentence is an example of a stimulus with a Topic facial expression (which should occur during the gloss "SWEET FOOD"): NEW RESTAURANT INCLUDE PASTA PIZZA SWEET FOOD MY SISTER COOK EXPERT. When the Topic face is perceived, then the stimulus has the approximate meaning: "The new restaurant has pasta and pizza. As for sweet foods (pastries), my sister is an expert chef." We have intentionally designed the stimulus so that it is performed at a human conversational speed without any long pauses during the signing that would emphasize the sentence boundary before "SWEET." This has been done so that the meaning of the stimulus is strongly affected by whether the viewer perceives the Topic facial expression. When the Topic face is not perceived, then the sentence boundary may be less clear (especially when the sentence is performed by an animated avatar that typically lacks the subtle acceleration and timing of a human signer). In such a case, the viewer may interpret "SWEET FOOD" as being the third item in the list of foods available at the restaurant; thereby the stimulus has the meaning: "The new restaurant has pasta, pizza, and sweet foods (pastries). My sister is an expert chef." One of the comprehension questions for this stimulus is: Does the new restaurant have sweet foods? The answer depends on whether the Topic facial expression was perceived and understood.

4.2 Example: WH-Word Questions

The following sentence is an example of a stimulus with a WH-Question facial expression (which should occur during the glosses "HER BIRTHDAY PARTY WHEN"): THAT MARY HER BIRTHDAY PARTY WHEN MARY DRUNK. When the WH-Question face is perceived, then the stimulus has the approximate meaning: "When is Mary's birthday party? Mary is drunk." When the WH-Question face is not perceived, then it may be less clear to the viewer where the sentence boundary is located. In such a case, the viewer may interpret "WHEN MARY DRUNK" as a question (albeit in English-like word order); thereby the stimulus would have the meaning: "It is Mary's birthday party. When did Mary got drunk?" One of the comprehension questions

for this stimulus is: Does Charlie know when the party is? (The signer appearing in the video is introduces as "Charlie" at the beginning of the study.) The participant is more likely to answer "no" to this question if the WH-Question facial expression was correctly perceived.

4.3 Example: Rhetorical Questions

The following sentence is an example of a stimulus with a RH-Question facial expression (which should occur during the glosses "WHY"): ALEX NOW GO-GO PARTIES WHY FINISH DIVORCE. When the RH-Question face is perceived, then the stimulus has the approximate meaning: "Alex is now often going to parties because he is divorced." When the RH-Question face is not perceived, then the sentence boundary may be less clear. In such a case, the viewer may interpret "WHY FINISH DIVORCE" as a question; thereby the stimulus has the meaning: "Alex is now often going to parties. Why did he get divorced?" One of the comprehension questions for this stimulus is: Does Charlie know why Alex started going to parties? The answer depends on whether the RH-Question facial expression was perceived and understood.

4.4 Example: Yes/No Questions

The following sentence is an example of a stimulus with a Yes/No Question facial expression (which should occur during the glosses "ALL FOOD CHEAP POINT"): BOB'S DINER THAT YOUR SISTER HER FAVORITE RESTAURANT ALL FOOD CHEAP POINT. When the YN-Question face is perceived, then the stimulus has the approximate meaning: "Bob's Diner is your sister's favourite restaurant. Is all the food cheap?" When the YN-Question face is not perceived, then the final sentence could appear to be a declarative statement. Thus, the stimulus has the meaning: "Bob's Diner is your sister's favourite restaurant. All the food is cheap." One of the comprehension questions for this stimulus is: Does Charlie know if the restaurant is expensive? If the YN-Question facial expression was correctly perceived and understood, then the participant is more likely to answer no to this question.

4.5 Example: Negation

The following sentence is an example of a stimulus with a Negation facial expression (which should occur during the glosses "HAVE SCIENCE CLASS"): ALEX TEND TAKE-UP MATH CLASS. NOW SEMESTER, SCHOOL HAVE SCIENCE CLASS. ALEX TAKE-UP TWO CLASS." When the Negation face is perceived, then the stimulus has the approximate meaning: "Alex usually takes math classes. This semester, the school doesn't have any science classes. Alex is taking two classes. " When the Negation face is not perceived, then the meaning of the middle sentence is inverted: "This semester, the school has science classes." One of the comprehension questions for this stimulus is: Does the school have science classes this semester? The answer depends on whether the Negation facial expression was perceived and understood.

4.6 Example: Emotional Affect

The following sentence is an example of a stimulus with an emotional affect facial expression (this example includes an angry facial expression during the entire sentence): LAST FRIDAY, MY BROTHER TAKE MY CAR. DRIVE SCHOOL. When the emotional affect facial expression is perceived, then the stimulus has the approximate meaning: "Last Friday, my brother took my car to drive to school." (The sentence has the subtext that the signer is upset about this.) When the emotional affect face is not perceived, then this subtext is not conveyed. One of the comprehension questions for this stimulus is: Is Charlie angry at his brother? The answer depends on whether the emotional facial expression was perceived and understood.

4.7 Likert-scale Questions

In addition to the four comprehension questions that are designed specifically for each stimulus, this collection also includes a set of Likert scale questions that can be used to measure participants' subjective evaluation of each. The set of Likert scale questions is identical for all of the stimuli, and it includes three subjective evaluation questions and four questions measuring whether participants' noticed a particular facial expression.

- "Good ASL grammar?": A subjective evaluation question of how grammatically correct was the presented signing with answers on a 1-to-10 Likert scale where 1 indicates bad and 10 perfect.
- "Easy to understand?": A subjective evaluation question on comprehensibility of the signed message with answers on a 1-to-10 scale where 1 indicates confusing and 10 clear.
- "Natural?": A subjective evaluation question on how naturally moving the signer appeared with answers on a 1-to-10 scale where 1 indicates that the signer moves like a robot and 10 that the signer moves like a person.
- "Did you notice a … facial expression?": Four questions in relation to how much participants noticed an emotional, negative, interrogative, or topic facial expression during the story with answers on a 1-to-10 scale from "yes" to "no".

The collection includes an HTML questionnaire with these Likert-scale questions and the options for the answers as radio buttons.

5. Facial Feature Extraction on Recordings

We used automatic face tracking software (Visage Technologies, 2014) to analyze the video recordings of the 48 ASL passages and produce files that contain information about the head pose and facial features of the human signer for each frame of the video. The tracking results, part of the collection, are shared as comma-separated values (CSV) files. Head pose data is given as translation from the camera in the 3 dimensions (x, y, z) and as head rotation (pitch, yaw, roll). The obtained facial features follow the MPEG-4 facial action parameters (Tekalp, 1999) for each frame of the video. For example, the eyebrow position in every frame is defined by 8 facial action parameters (FAP30-FAP37) as the vertical and horizontal displacement of the left and right eyebrow from a neutral pose of the signer's face. This information could be used by future researchers to

animate the face of a virtual human character (Pandzic and Forchheimer, 2003) performing these stimuli passages. Such a character could be displayed as a baseline for comparison in an experimental evaluation study.

For optimal results, the Visage software was used in offline mode. The quality of the results is bounded by the performance of the software on the video recordings and the initial manual process of mask fitting to the face as shown in Fig. 2. For example, the tracker may lose the face if the head movement is too fast or if large parts of the face are covered, e.g. by the hands. We observed that this is happening for 0%-7.6% (avg. 1.6%) of the story duration in our stimuli collection. In this case, the lost frames are indicated with a tracking status other than "OK" in the comma-separated values file, and all the extracted head and facial features would normally have the value 0 in such cases. We processed the data and filled in the values of the lost frames using spline interpolation (smoothing degree 1) while maintaining the tracking status information. Although interpolation may work well for the facial feature values, it can sometimes be problematic for head rotation, because it is currently represented in the form of Euler angles (pitch, yaw, roll). We advise future researchers to consider first converting the head rotation into another representation (e.g. quaternions) and then to apply interpolation techniques to fill in the rotation values for the lost frames.

Figure 2: Fitted face shape mask in Visage software.

6. Stimuli Quality as Measured by Participants in Previous Studies

This stimuli collection contains passages appropriate for use during a user study evaluating facial expressions in ASL animations. A subset of these passages and comprehension questions has already been used in prior studies at our laboratory (Kacorri, Lu, and Huenerfauth, 2013a; Kacorri, Harper, and Huenerfauth, 2013). The following stimuli in this collection were included in these two prior studies: E1, E2, E4, E5, E6, E8, W2, W3, W4, Y3, Y5, Y6, R3, R5, R7, N1, N2, N3, T3, T4, T5.

The first study consisted of a user study in which native ASL signers viewed human videos (with natural facial expressions) and ASL animations (without any facial expressions) and responded to comprehension questions (Kacorri, Lu, and Huenerfauth, 2013a). In the second study, identical stimuli were shown and similar participants were recruited, but in this study, the participants viewed the animations on a computer screen that was mounted above a desktop eye-tracking system that tracked their gaze location on the stimulus (Kacorri, Harper, and Huenerfauth, 2013). Full details of the studies appear in the original publications. Figure 3 presents the human video and the no-facial-expression animation results from these two studies. Bars are shown separately for each category of stimuli: emotional affect, negation, topic, WH-question, YN-question, and RH-question. Here we see that the stimuli with facial expressions received higher comprehension question scores than the stimuli without facial expressions, which suggests the suitability of these questions for user studies evaluating the perception of facial expressions. In future work, we intend to conduct more rigorous studies of the efficacy of these stimuli and questions, and we intend to examine the quality of the additional stimuli that were not included in these two prior studies. We also welcome feedback and improvements to the stimuli from other researchers who make use of this collection.

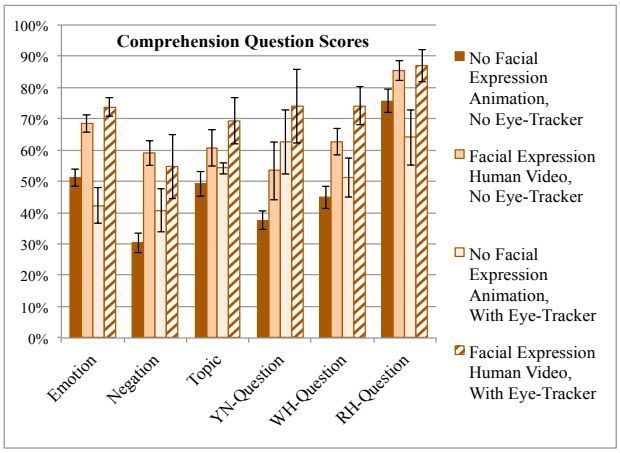

Figure 3: Comprehension question scores from the subset of stimuli in the collection used in prior evaluation studies.

7. Availability of the Collection

As with prior ASL corpora resources released by our laboratory (Lu and Huenerfauth, 2009; Lu and Huenerfauth, 2012), this stimuli collection is available for use by other sign language animation researchers, details appear here: http://latlab.cs.qc.cuny.edu/lrec2014

We invite members of the research community to provide feedback to us about the stimuli in this collection, and we welcome recommendations of additional stimuli designs or edits that would enhance the collection (which we would look forward to incorporating into a future release of this resource). While the current collection of stimuli has not yet been rigorously evaluated, we see a benefit for rapidly releasing this resource to the research community for use and feedback. Ultimately, the field of sign language animation synthesis may benefit from the community identifying a standard set of evaluation stimuli and questions for system evaluation, to better enable comparison of systems and progress in the field.

In future work at our laboratory, we are continuing to investigate the design of animation models for ASL facial expressions, and we are continuing to make use of these stimuli and questions to evaluate the quality of our animation results.

8. Acknowledgements

This material is based upon work supported in part by the US. National Science Foundation under award number 0746556 and 1065009. This work has also been supported by Visage Technologies AB through a free academic license for character animation software. Miriam Morrow, Jonathan Lamberton, and Jennifer Marfino assisted with the design of stimuli, and we are grateful for advice and feedback from Susan Fischer.

9. References

Gibet, S., Courty, N., Duarte, K., and Naour, T. L. 2011. The SignCom system for data-driven animation of interactive virtual signers: Methodology and Evaluation. ACM Transactions on Interactive Intelligent Systems (TiiS), 1(1), 6.

Huenerfauth, M. 2004a. American Sign Language Spatial Representations for an Accessible User-Interface. In Proc. of the 3rd International Conference on Universal Access in Human-Computer Interaction. Las Vegas, NV, USA.

Huenerfauth, M. 2004b. Spatial and planning models of ASL classifier predicates for ma- chine translation. In 10th International Conference on Theoretical and Methodological Is-sues in Machine Translation: TMI.

Huenerfauth, M., Zhao, L., Gu, E., and Allbeck, J. 2007. Evaluating American Sign Language Generation through the Participation of Native ASL Signers. In Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2007), Tempe, Arizona, USA.

Huenerfauth, M. 2008. Evaluation of a psycholinguistically motivated timing model for animations of American Sign Language. In Proc. 10th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2008).

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

Kacorri, H. 2013. Models of linguistic facial expressions for American Sign Language animation. ACM SIGACCESS Accessibility and Computing, (105), 19-23.

Kacorri, H., Lu, P., and Huenerfauth, M. 2013a. Effect of Displaying Human Videos During an Evaluation Study of American Sign Language Animation. ACM Trans. Access. Comput. 5, 2, Article 4.

Kacorri, H., Lu, P., and Huenerfauth, M. 2013b. Evaluating Facial Expressions in American Sign Language Animations for Accessible Online Information. In Proc. International Conference on Universal Access in Human Computer Interaction (UAHCI). Las Vegas, NV, USA.

Kacorri, H., Harper, A., and Huenerfauth, M. 2013. Comparing native signers' perception of American Sign Language animations and videos via eye tracking. In Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility (p. 9). ACM.

Liu, J., Liu, B., Zhang, S., Yang, F., Yang, P., Metaxas, D.N., and Neidle, C. 2013. Recognizing Eyebrow and Periodic Head Gestures Using CRFs for Non-Manual Grammatical Marker Detection in ASL. In Proceedings of FG 2013: 10th IEEE International Conference on Automatic Face and Gesture Recognition, Shanghai China.

Lu, P., and Huenerfauth, M. 2009. Accessible Motion-Capture Glove Calibration Protocol for Recording Sign Language Data from Deaf Subjects. In Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2009), Pittsburgh, Pennsylvania, USA.

Lu, P., and Huenerfauth, M. 2012. CUNY American Sign Language Motion-Capture Corpus: First Release. In Proceedings of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions between Corpus and Lexicon, The 8th International Conference on Language Resources and Evaluation (LREC 2012), Istanbul, Turkey.

Mitchell, R., Young, T., Bachleda, B., and Karchmer, M. 2006. How many people use ASL in the United States? Why estimates need updating. Sign Lang Studies, 6(3):306-335.

Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., Lee, R. 2000. *The Syntax of American Sign Language: functional categories and hierarchical structure.* Cambridge, MA: MIT Press.

Pandzic, I. S., and Forchheimer, R. (eds.). 2003. MPEG-4 facial animation: the standard, implementation and applications. Wiley.

Schmidt, C., Koller, O., Ney, H., Hoyoux, T., and Piater, J. 2013. Enhancing Gloss-Based Corpora with Facial Features Using Active Appearance Models. Third International Symposium on Sign Language Translation and Avatar Technology (SLTAT).

Schnepp, J., Wolfe, R., McDonald, J. C., Jorge Toro. 2012. Combining emotion and facial nonmanual signals in synthesized american sign language. In Proc. 14th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2012), 249-250.

Tekalp, M. 1999. Face and 2-D Mesh Animation in MPEG-4, Tutorial Issue on the MPEG-4 Standard, Image Communication Journal, Elsevier.

Traxler, C. 2000. The Stanford achievement test, 9th edition: national norming and performance standards for deaf and hard-of-hearing students. J Deaf Stud & Deaf Educ, 5:4, pp. 337-348.

Visage Technologies. 2014. Visage Technologies Face Tracking. Retrieved February 4, 2014, from http://www.visagetechnologies.com/products/visagesdk/facetrack/.

Wolfe, R., Cook, P., McDonald, J. C., and Schnepp, J. 2011. Linguistics as structure in computer animation: Toward a more effective synthesis of brow motion in American Sign Language. Sign Language & Linguistics, 14(1), 179-199.

Yu, X., Huang, J., Zhang, S., Yan, W., and Metaxas, D. N. 2013. Pose-free Facial Landmark Fitting via Optimized Part Mixtures and Cascaded Deformable Shape Model. In IEEE International Conference on Computer Vision.