Design and Evaluation of an American Sign Language Generator

Matt Huenerfauth
Computer Science Department
CUNY Queens College
The City University of New York
65-30 Kissena Boulevard
Flushing, NY 11375 USA
matt@cs.qc.cuny.edu

Liming Zhao, Erdan Gu, Jan Allbeck
Center for Human Modeling & Simulation
University of Pennsylvania
3401 Walnut Street
Philadelphia, PA 19104 USA
{liming,erdan,allbeck}
@seas.upenn.edu

Abstract

We describe the implementation and evaluation of a prototype American Sign Language (ASL) generation component that produces animations of ASL classifier predicates, some frequent and complex spatial phenomena in ASL that no previous generation system has produced. We discuss some challenges in evaluating ASL systems and present the results of a user-based evaluation study of our system.

1 Background and Motivations

American Sign Language (ASL) is a natural language with a linguistic structure distinct from English used as a primary means of communication for approximately one half million people in the U.S. (Mitchell et al., 2006). A majority of deaf 18-year-olds in the U.S. have an English reading level below that of an average 10-year-old hearing student (Holt, 1991), and so software to translate English text into ASL animations can improve many people's access to information, communication, and services. Previous English-to-ASL machine translation projects (Sáfár & Marshall, 2001; Zhou et al., 2000) could not generate classifier predicates (CPs), phenomena in which signers use special hand movements to indicate the location and movement of invisible objects in space around them (representing entities under discussion). Because CPs are frequent in ASL and necessary for conveying many concepts, we have developed a CP generator that can be incorporated into a full English-to-ASL machine translation system.

During a CP, signers use their hands to position, move, trace, or re-orient imaginary objects in the space in front of them to indicate the location, movement, shape, contour, physical dimension, or some other property of corresponding real world entities under discussion. CPs consist of a semantically meaningful handshape and a 3D hand movement path. A handshape is chosen from a closed set based on characteristics of the entity described (whether it be a vehicle, human, animal, etc.) and what aspect of the entity the signer is describing (surface, position, motion, etc). For example, the sentence "the car parked between the cat and the house" could be expressed in ASL using 3 CPs. First, a signer performs the ASL sign HOUSE while raising her eyebrows (to introduce a new entity as a topic). Then, she moves her hand in a "Spread C" handshape (Figure 1) forward to a point in space where a miniature house could be envisioned. Next, the signer performs the sign CAT with eyebrows raised and makes a similar motion with a "Hooked V" handshape to a location where a cat could be imagined. Finally, she performs the sign CAR (with eyebrows raised) and uses a "Number 3" handshape to trace a path that stops at between the 'house' and the 'cat.' Her other hand makes a flat surface for the 'car' to park on. (Figure 3 will show our system's animation.)

Figure 1: ASL handshapes: Spread C (bulky object), Number 3 (vehicle), Hooked V (animal), Flat (surface).

Proceedings of the Workshop on Embodied Language Processing, pages 51–58,
Prague, Czech Republic, June 28, 2007.

2 System Design and Implementation

We have built a prototype ASL generation module that could be incorporated into an English-to-ASL machine translation system. When given a 3D model of the arrangement of a set of objects whose location and movement should be described in ASL, our system produces an animation of ASL sentences containing classifier predicates to describe the scene. Classifier predicates are the way such spatial information is typically conveyed in ASL. Since this is the first ASL generation system to produce classifier predicate sentences (Huenerfauth, 2006b), we have also conducted an evaluation study in which native ASL signers compared our system's animations to the current state of the art: Signed English animations (described later).

2.1 Modeling the Use of Space

To produce classifier predicates and other ASL expressions that associate locations in space around a signer with entities under discussion, an English-to-ASL system must model what objects are being discussed in an English text, and it must map placeholders for these objects to locations in space around the signer's body. The input to our ASL classifier predicate generator is an explicit 3D model of how a set of placeholders representing discourse entities are positioned in the space around the signing character's body (Huenerfauth, 2006b). This 3D model is "mapped" onto a volume of space in front of the signer's torso, and this model is used to guide the motion of the ASL signer's hands during the performance of classifier predicates describing the motion of these objects

The model encodes the 3D location (center-of-mass) and orientation values of the set of objects that we want to our system describe using ASL animation. For instance, to generate the "car parking between the cat and the house" example, we would pass our system a model with three sets of location (x, y, z coordinates) and orientation (x, y, z, rotation angles) values: for the cat, the car, and the house. Each 3D placeholder also includes a set of bits that represent the set of possible ASL classifier handshapes that can be used to describe it.

While this 3D model is given as input to our prototype classifier predicate generator, when part of a full generation system, virtual reality "scene visualization" software can be used to produce a 3D model of the arrangement and movement of objects discussed in an English input text (Badler et al., 2000; Coyne and Sproat, 2001).

2.2 Template-Based Planning Generation

Given the 3D model above, the system uses a planning-based approach to determine how to move the signer's hands, head-tilt, and eye-gaze to produce an animation of a classifier predicate. The system stores a library of templates representing the various kinds of classifier predicates it may produce. These templates are planning operators (they have logical pre-conditions, monitored termination conditions, and effects), allowing the system to trigger other elements of ASL signing performance that may be required during a grammatically correct classifier predicate (Huenerfauth, 2006b). Each planning operator is parameterized on an object in the 3D model (and its 3D coordinates); for instance, there is a templated planning operator for generating an ASL classifier predicate to show a "parking" event. The specific location/orientation of the vehicle that is parking would be the parameter passed to the planning operator.

There is debate in the ASL linguistics community about the underlying structure of classifier predicates and the generation process by which signers produce them. Our parameterized template approach mirrors one recent linguistic model (Liddell, 2003), and the implementation and evaluation of our prototype generator will help determine whether this was a good choice for our system.

2.3 Multi-Channel Syntax Representation

While strings and syntax trees are used to represent written languages inside of NLP software, these encodings are difficult to adapt to a sign language. ASL lacks a standard writing system, and the multichannel nature of an ASL performance makes it difficult to encode in a linear single-channel string. This project developed a new formalism for representing a linguistic signal in a multi-channel manner and for encoding temporal coordination and non-coordination relationships between portions of the signal (Huenerfauth, 2006a). The output of our planner is a tree-like structure that represents the animation to be synthesized. The tree has two kinds of non-terminal nodes: some indicate that their children should be performed in sequence (like a traditional linguistic syntax tree), and some non-terminals indicate that their children should be performed in parallel (e.g. one child subtree may specify the movement of the arms, and another, the facial expression). In this way, the structure can encode how multiple parts of the sign language performance should be coordinated over time while still leaving flexibility to the exact timing of events – see Figure 2. In earlier work, we have argued that this representation is sufficient for encoding ASL animation (Huenerfauth, 2006a), and the implementation and evaluation of our system (using this formalism) will help test this claim.

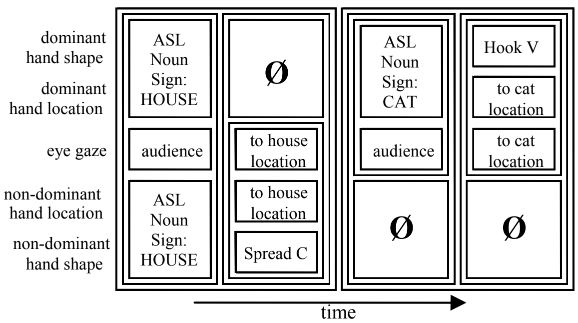

Figure 2: A multichannel representation for the sentence "The cat is next to the house." This example shows handshape, hand location, and eye gaze direction – some details omitted from the example: hand orientation, head tilt, and brow-raising. Changes in timing of individual animation events causes the structure to stretch in the time dimension (like an HTML table).

2.4 Creating Virtual Human Animation

After planning, the system has a tree-structure that specifies activities for parts of the signer's body. Non-terminal nodes indicate whether their children are performed in sequence or in parallel, and the terminal nodes (the inner rectangles in Figure 2) specify animation events for a part of the signer's body. Nodes' time durations are not yet specified (since the human animation component would know the time that movements require, not the linguistic planner). So, the generator queries the human animation system to calculate an estimated time duration for each body animation event (each terminal node), and the structure is then 'balanced' so that if several events are meant to occur in parallel, then the shorter events are 'stretched out.' (The linguistic system can set max/min times for some events prior to the animation processing.)

2.5 Eye-Gaze and Brow Control

The facial model is implemented using the Greta facial animation engine (Pasquariello and Pelachaud, 2001). Our model controls the motion of the signer's eye-brows, which can be placed in a "raised" or "flat" position. The eye motor control repertoire contains three behaviors: fixation on a 3D location in space around the signer's body, smooth pursuit of a moving 3D location, and eye-blinking. Gaze direction is computed from the location values specified inside the 3D model, and the velocity and time duration of the movement are determined by the timing values inside the tree-structure output from the planner. The signer's head tilt changes to accommodate horizontal or vertical gaze shifts greater than a set threshold. When performing a "fixation" or "smooth pursuit" with the eye-gaze, the rate of eye blinking is decreased. Whenever the signer's eye-gaze is not otherwise specified for the animation performance, the default behavior is to look at the audience.

2.6 Planning Arm Movement

Given the tree-structure with animation events, the output of arm-planning should be a list of animation frames that completely specify the rotation angles of the joints of the signer's hands and arms. The hand is specified using 20 rotation angles for the finger joints, and the arm is specified using 9 rotation angles: 2 for the clavicle joint, 3 for the shoulder joint, 1 for the elbow joint, and 3 for the wrist. The linguistic planner specifies the handshape that should be used for specific classifier predicates; however, the tree-structure specifies the arm movements by giving a target location for the center of the signer's palm and a target orientation value for the palm. The system must find a set of clavicle, shoulder, elbow, and wrist angles that get the hand to this desired location and palm orientation. In addition to reaching this target, the arm pose for each animation frame must be as natural as possible, and the animation between frames must be smooth. The system uses an inverse kinematics (IK) which automatically favors natural arm poses. Using the wrist as the end-effector, an elbow angle is selected based on the distance from shoulder to the target, and this elbow angle is fixed. We next compute a set of possible shoulder and wrist rotation angles in order to align the signer's hand with the target palm orientation. Disregarding elbow angles that force impossible wrist joint angles, we select the arm pose that is collision free and is the most natural, according to a shoulder strength model (Zhao et al., 2005).

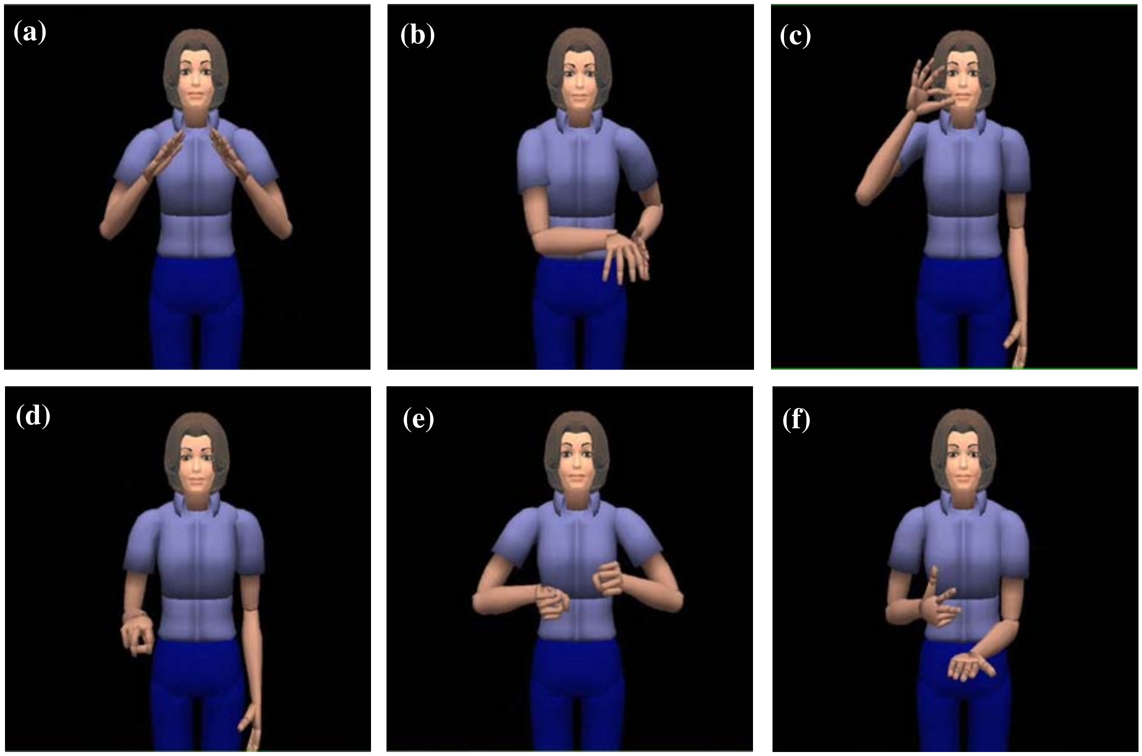

Figure 3: Images from our system's animation of a classifier predicate for "the car parked between the house and the cat." (a) ASL sign HOUSE, eyes at audience, brows raised; (b) Spread C handshape and eye gaze to house location; (c) ASL sign CAT, eyes at audience, brows raised; (d) Hooked V handshape and eye gaze to cat location; (e) ASL sign CAR, eyes at audience, brows raised; (f) Number 3 handshape (for the car) parks atop Flat handshape while the eye gaze tracks the movement path of the car.

2.7 Synthesizing Virtual Human Animation

This animation specification is performed by an animated human character in the Virtual Human Testbed (Badler et al., 2005). Because the Greta system used a female head with light skin tone, a female human body was chosen with matching skin. The character was dressed in a blue shirt and pants that contrasted with its skin tone. To make the character appear to be a conversational partner, the "camera" inside the virtual environment was set at eye-level with the character and at an appropriate distance for ASL conversation.

2.8 Coverage of the Prototype System

Our prototype system can be used to translate a limited range of English sentences (discussing the locations and movements of a small set of people or objects) into animations of an onscreen human-like character performing ASL classifier predicates to convey the locations and movements of the entities in the English text. Table 1 includes shorthand transcripts of some ASL sentence animations produced by the system; the first sentence corresponds to the classifier predicate animation in Figure 3.

3 Issues in Evaluating ASL Generation

There has been little work on developing evaluation methodologies for sign language generation or MT systems. Some have shown how automatic string-based evaluation metrics fail to identify correct sign language translations (Morrisey and Way, 2006), and they propose building large parallel written/sign corpora containing more syntactic and semantic information (to enable more sophisticated metrics to be created). Aside from the expense of creating such corpora, we feel that there are several factors that motivate user-based evaluation studies for sign language generation systems – especially for those systems that produce classifier predicates. These factors include some unique linguistic properties of sign languages and the lack of standard writing systems for most sign languages, like ASL.

Most automatic evaluation approaches for generation or MT systems compare a string produced by a system to a human-produced "gold-standard" string. Sign languages usually lack written forms that are commonly used or known among signers. While we could invent an artificial ASL writing system for the generator to produce as output (for evaluation purposes only), it's not clear that human ASL signers could accurately or consistently produce written forms of ASL sentences to serve as "gold standards" for such an evaluation. Further, real users of the system would never be shown artificial written ASL; they would see animation output. Thus, evaluations based on strings would not test the full process – including the synthesis of the "string" into an animation – when errors may arise.

Another reason why string-based evaluation metrics are not well-suited to ASL is that sign languages have linguistic properties that can confound string-edit-distance-based metrics. ASL consists of the coordinated movement of several parts of the body in parallel (i.e. face, eyes, head, hands), and so a string listing the set of signs performed is a lossy representation of the original performance (Huenerfauth, 2006a). The string may not encode the non-manual parts of the sentence, and so string-based metrics would fail to consider those important aspects. Discourse factors (e.g. topicalization) can also result in movement phenomena in ASL that may scramble the sequence of signs in the sentence without substantially changing its semantics; such movement would affect string-based metrics significantly though the sentence meaning may change little. The use of head-tilt and eye-gaze during the performance of ASL verb signs may also license the dropping of entire sentence constituents (Neidle et. al, 2000). The entities discussed are associated with locations in space around the signer at which head-tilt or eye-gaze is aimed, and thus the constituent is actually still expressed although no manual signs are performed for it. Thus, an automatic metric may penalize such a sentence (for missing a constituent) while the information is still there. Finally, ASL classifier predicates convey a lot of information in a single complex 'sign' (handshape indicates semantic category, movement shows 3D path/rotation), and it is unclear how we could "write" the 3D data of a classifier predicate in a string-based encoding or how to calculate an edit-distance between a 'gold standard' classifier predicate and a generated one.

4 Evaluation of the System

We used a user-based evaluation methodology in which human native ASL signers are shown the output of our generator and asked to rate each animation on ten-point scales for understandability, naturalness of movement, and ASL grammatical correctness. To evaluate whether the animation conveyed the proper semantics, signers were also asked to complete a matching task. After viewing a classifier predicate animation produced by the system, signers were shown three short animations showing the movement or location of the set of objects that were described by the classifier predicate. The movement of the objects in each animation was slightly different, and signers were asked to select which of the three animations depicted the scene that was described by the classifier predicate.

Since this prototype is the first generator to produce animations of ASL classifier predicates, there are no other systems to compare it to in our study. To create a lower baseline for comparison, we wanted a set of animations that reflect the current state of the art in broad-coverage English-to-sign

English Gloss	**ASL Sentence with Classifier Predicates (CPs)**	**Signed English Sentence**
The car parks between the house and the cat.	ASL sign HOUSE; CP: house location; sign CAT; CP: cat location; sign CAR; CP: car path.	THE CAR PARK BETWEEN THE HOUSE AND THE CAT
The man walks next to the woman.	ASL sign WOMAN; CP: woman location; sign MAN; CP: man path.	THE MAN WALK NEXT-TO THE WOMAN
The car turns left.	ASL sign CAR; CP: car path.	THE CAR TURN LEFT
The lamp is on the table.	ASL sign TABLE; CP: table location; sign LIGHT; CP: lamp location.	THE LIGHT IS ON THE TABLE
The tree is near the tent.	ASL sign TENT; CP: tent location; sign TREE; CP: tree location.	THE TREE IS NEAR THE TENT
The man walks between the tent and the frog.	ASL sign TENT; CP: tent location; sign FROG; CP: frog location; sign MAN; CP: man path.	THE MAN WALK BETWEEN THE TENT AND THE FROG
The man walks away from the woman.	ASL sign WOMAN: CP: woman location; sign MAN; CP: man path.	THE MAN WALK FROM THE WOMAN
The car drives up to the house.	ASL sign HOUSE; CP: house location; sign CAR; CP: car path.	THE CAR DRIVE TO THE HOUSE
The man walks up to the woman.	ASL sign WOMAN; CP: woman location; sign MAN; CP: man path.	THE MAN WALK TO THE WOMAN
The woman stands next to the table.	ASL sign TABLE; CP: table location; sign WOMAN; CP: woman location.	THE WOMAN STAND NEXT-TO THE TABLE

Table 1: ASL and Signed English sentences included in the evaluation study (with English glosses).

translation. Since there are no broad-coverage English-to-*ASL* MT systems, we used Signed English transliterations as our lower baseline. Signed English is a form of communication in which each word of an English sentence is replaced with a corresponding sign, and the sentence is presented in original English word order without any accompanying ASL linguistic features such as meaningful facial expressions or eye-gaze.

Ten ASL animations (generated by our system) were selected for inclusion in this study based on some desired criteria. The ASL animations consist of classifier predicates of movement and location – the focus of our research. The categories of people and objects discussed in the sentences require a variety of ASL handshapes to be used. Some sentences describe the location of objects, and others describe movement. The sentences describe from one to three objects in a scene, and some pairs of sentences actually discuss the same set of objects, but moving in different ways. Since the creation of a referring expression generator was not a focus of our prototype, all referring expressions in the animations are simply an ASL noun phrase consisting of a single sign – some one-handed and some two-handed. Table 1 lists the ten classifier predicate animations we selected (with English glosses).

For the "matching task" portion of the study, three animated visualizations were created for each sentence showing how the objects mentioned in the sentence move in 3D. One animation was an accurate visualization of the location/movement of the objects, and the other two animations were "confusables" – showing orientations/movements for the objects that did not match the classifier predicate animations. Because we wanted to evaluate the classifier predicates (and not the referring expressions), the set of objects that appeared in all three visualizations for a sentence was the same. Thus, it was the movement and orientation information conveyed by the classifier predicate (and not the object identity conveyed by the referring expression) that would distinguish the correct visualization from the confusables. For example, the following three visualizations were created for the sentence "the car parks between the cat and the house" (the cat and house remain in the same location in each): (1) a car drives on a curved path and parks at a location between a house and a cat, (2) a car drives between a house and a cat but continues driving past them off camera, and (3) a car starts at a location between a house and a cat and drives to a location that is not between them anymore.

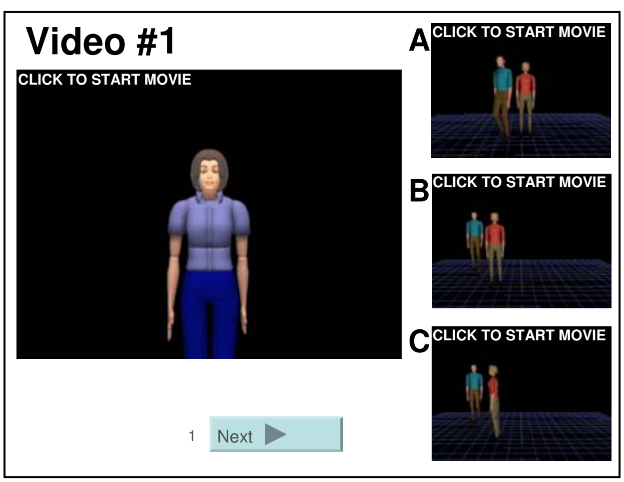

Figure 4: Screenshot from evaluation program.

To create the Signed English animations for each sentence, some additional signs were added to the generator's library of signs. (ASL does not traditionally use signs such as "THE" that are used in Signed English.) A sequence of signs for each Signed English transliteration was concatenated, and the synthesis sub-component of our system was used to calculate smooth transitional movements for the arms and hands between each sign in the sentence. The glosses for the ten Signed English transliterations are also listed in Table 1.

4.1 User-Interface for Evaluation Study

An interactive slideshow was created with one slide for each of the 20 animations (10 from our ASL system, 10 Signed English). On each slide, the signing animation was shown on the left of the screen, and the three possible visualizations of that sentence were shown to the right (see Figure 4). The slides were placed in a random order for each of the participants in the study. A user could replay the animations as many times as desired before going to the next signing animation. Subjects were asked to rate each of these animations on a 1-to-10-point scale for ASL grammatical correctness, understandability, and naturalness of movement. Subjects were also asked to select which of the three animated visualizations (choice "A," "B," or "C") matched the scene as described in the sentence performed by the virtual character.

After these slides, 3 more slides appeared containing animations from our generator. (These were repeats of 3 animations used in the main part of the study.) These three slides only showed the "correct" animated visualization for that sentence. For these last three slides, subjects were instead asked to comment on the animation's speed, colors/lighting, hand visibility, correctness of hand movement, facial expression, and eye-gaze. Signers were also asked to write any comments they had about how the animation should be improved.

4.2 Recruitment and Screening of Subjects

Subjects were recruited through personal contacts in the deaf community who helped identify friends, family, and associates who met the screening criteria. Participants had to be native ASL signers – many deaf individuals are non-native signers who learned ASL later in life (and may accept English-like signing as being grammatical ASL). Subjects were preferred who had learned ASL since birth, had deaf parents that used ASL at home, and/or attending a residential school for the deaf as a child (where they were immersed in an ASL-signing community). Of our 15 subjects, 8 met all three criteria, 2 met two criteria, and 5 met one (1 grew up with ASL-signing deaf parents and 4 attended a residential school for the deaf from an early age).

During the study, instructions were given to participants in ASL, and a native signer was present during 13 of the 15 sessions to answer questions or to explain experimental procedures. This signer engaged the participants in conversation in ASL before the session to produce an ASL-immersive environment. Participants were given instructions in ASL about how to score each category. For grammaticality, they were told that "perfect ASL grammar" would be a 10, but "mixed-up" or "English-like" grammar should be a 1. For understandability, "easy to understand" sentences should be a 10, but "confusing" sentences should be a 1. For naturalness, animations in which the signer moved "smoothly, like a real person" should be a 10, but animations in which the signer moved in a "choppy" manner "like a robot" should be a 1.

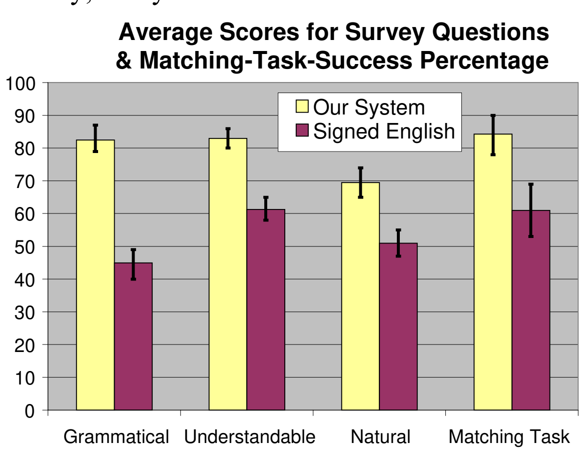

Figure 5: Grammaticality, understandability, naturalness, and matching-task-success scores.

4.3 Results of the Evaluation

Figure 5 shows average scores for grammaticality, understandability, naturalness, and matching-task-success percentage for the animations from our system compared to the Signed English animations. Our system's higher scores in all categories is significant ($\alpha = 0.05$, pairwise Mann-Whitney U tests with Bonferonni-corrected p-values).

Subjects were asked to comment on the animation speed, color, lighting, visibility of the hands, correctness of hand movement, correctness of facial expressions, correctness of eye-gaze, and other ways of improving the animations. Of the 15 subjects, eight said that some animations were a little slow, and one felt some were very slow. Eight subjects wanted the animations to have more facial expressions, and 4 of these specifically mentioned nose and mouth movements. Four subjects said the signer's body should seem more loose/relaxed or that it should move more. Two subjects wanted the signer to show more emotion. Two subjects felt that eye-brows should go higher when raised, and three felt there should be more eye-gaze movements. Two subjects felt the blue color of the signer's shirt was a little too bright, and one disliked the black background. Some subjects commented on particular ASL signs that they felt were performed incorrectly. For example, three discussed the sign "FROG": one felt it should be performed a little more to the right of its current location, and another felt that the hand should be oriented with the fingers aimed more to the front. Some participants commented on the classifier predicate portions of the performance. For example, in the sentence "the car parked between the cat and the house," one subject felt it would be better to use the non-dominant hand to hold the location of the house during the car's movement instead of using the non-dominant hand to create a platform for the dominant hand (the car) to park upon.

5 Conclusions and Future Work

Unlike an evaluation of a broad-coverage NLP system, during which we obtain performance statistics for the system as it carries out a linguistic task on a large corpus or "test set," this paper has described an evaluation of a prototype system. We were not measuring the linguistic coverage of the system but rather its functionality. Did signers agree that the animation output: (1) is actually a grammatically-correct and understandable classifier predicate and (2) conveys the information about the movement of objects in the 3D scene being described? We expected to find animation details that could be improved in future work; however, since there are currently no other systems capable of generating ASL classifier predicate animations, any system receiving an answer of "yes" to questions (1) and (2) above is an improvement to the state of the art.

Another contribution of this initial evaluation is that it serves as a pilot study to help us determine how to better evaluate sign language generation systems in the future. We found that subjects were comfortable critiquing ASL animations, and most suggested specific (and often subtle) elements of the animation to be improved. Their feedback suggested new modifications we can make to the system (and then evaluate again in future studies). Because subjects gave such high quality feedback, future studies will also elicit such comments.

During the study, we also experimented with recording a native ASL signer (using a motion-capture suit and datagloves) performing classifier predicates. We tried to use this motion-capture data to animate a virtual human character superficially identical to the one used by our system. We hoped that this character controlled by human movements could serve as an upper-baseline in the evaluation study. Unfortunately, the motion-capture data we collected contained minor errors that required post-processing clean-up, and the resulting animations contained enough movement inaccuracies that native ASL signers who viewed them felt they were actually less understandable than our system's animations. In future work, we intend to explore alternative upper-baselines to compare our system's animations to: animation from alternative motion-capture techniques, hand-coded animations based on a human's performance, or simply a video of a human signer performing ASL sentences.

Acknowledgements

National Science Foundation Award #0520798 "SGER: Generating Animations of American Sign Language Classifier Predicates" (Universal Access, 2005) supported this work. Software was donated by UGS Tecnomatix and Autodesk. Thank you to Mitch Marcus, Martha Palmer, and Norman Badler.

References

N.I. Badler, J. Allbeck, S.J. Lee, R.J. Rabbitz, T.T. Broderick, and K.M. Mulkern. 2005. New behavioral paradigms for virtual human models. SAE Digital Human Modeling.

N. Badler, R. Bindiganavale, J. Allbeck, W. Schuler, L. Zhao, S. Lee, H. Shin, & M. Palmer. 2000. Parameterized action representation & natural language instructions for dynamic behavior modification of embodied agents. AAAI Spring.

R. Coyne and R. Sproat. 2001. WordsEye: an automatic text-to-scene conversion system. ACM SIGGRAPH.

J.A. Holt. 1991. Demographic, Stanford Achievement Test - 8th Edition for Deaf and Hard of Hearing Students: Reading Comprehension Subgroup Results.

É . Sáfár & I. Marshall. 2001. The architecture of an English-text-to-Sign-Languages translation system. Recent Advances in Natural Language Processing.

Matt Huenerfauth. In Press. Representing American Sign Language classifier predicates using spatially parameterized planning templates. In M. Banich and D. Caccamise (Eds.), *Generalization.* Mahwah: LEA.

Matt Huenerfauth. 2006a. Representing Coordination and Non-Coordination in American Sign Language Animations. *Behaviour & Info. Technology*, 25:4.

Matt Huenerfauth. 2006b. Generating American Sign Language Classifier Predicates for English-to-ASL Machine Translation. Dissertation, U. Pennsylvania.

Liddell, S. 2003. Grammar, Gesture, and Meaning in American Sign Language. UK: Cambridge U. Press.

R.E. Mitchell, T.A. Young, B. Bachleda, & M.A. Karchmer. 2006. How Many People Use ASL in the United States? Why estimates need updating. *Sign Language Studies*, 6:3.

S. Morrissey & A. Way. 2006. Lost in Translation: The problems of using mainstream MT evaluation metrics for sign language translation. *5th SALTMIL Workshop on Minority Languages, LREC-200*

C. Neidle, J. Kegl, D. MacLaughlin, B. Bahan, & R.G. Lee. 2000. *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure.* Cambridge: MIT.

S. Pasquariello & C. Pelachaud. 2001. Greta: A simple facial animation engine. In 6th Online World Conference on Soft Computing in Industrial Applications.

L. Zhao, K. Kipper, W. Schuler, C. Vogler, N.I. Badler, & M. Palmer. 2000. Machine Translation System from English to American Sign Language. Assoc. for MT in the Americas.

L. Zhao, Y. Liu, N.I. Badler. 2005. Applying empirical data on upper torso movement to real-time collision-free reach tasks. SAE Digital Human Modeling.