Machine Translation (2021) 35:345–362
https://doi.org/10.1007/s10590-021-09279-9

Generalizing a model for animating adverbs of manner in American Sign Language

Robyn Moncrief[1]

Received: 22 June 2020 / Accepted: 16 July 2021 / Published online: 24 August 2021

Abstract
This work aims to show that a model produced to generate adverbs of manner can be generalized and applied to a variety of neutral animated signs for avatar sign language synthesis. This paper presents the generalization of a new approach that was first presented at SLTAT 2019 in Hamburg for modeling language processes that manifest themselves as modifications to the visual-manual channel. This work discusses extensions for generalizability to the model to be effective for a broader range of signs including one-handed and two-handed signs, repeating and non-repeating signs, signs with contact, and additional rotational adjustments to the wrists. This paper also includes interim results from an ongoing user study.

Keywords Avatar technology · Sign synthesis · Adverbs in ASL · Motion planning

1 Introduction

The quality of animated sign translation systems hinges on the signing avatar. Signed utterances must be grammatically correct and easily understood. Although an avatar's appearance is important to its legibility, its motion has an even greater impact (Malala et al. 2018). The computation of an avatar's movement is key to sign synthesis. It is determined by a combination of:

- Path computation, related to motion planning (Barraquand and Latombe 1991),
- Timing along that path
- The determination of joint participation in creating the path (Mcdonald et al. 2016, 2005)
- Ancillary motions required to support the clarity of the utterance (Schnepp et al. 2012)

✉ Robyn Moncrief
rmoncrie@cdm.depaul.edu

1 College of Computing and Digital Media, CDM 822, DePaul University, 243 S. Wabash Ave, Chicago, IL 60604, USA

This case study presents the generalization of a new approach to modeling sign language processes that manifest themselves as modifications to the visual-manual channel.

2 Adverbs of manner

Automatic translation between spoken and signed languages presents many challenges. As American Sign Language (ASL) is an independent natural language (Stokoe 1960), there is no one-to-one mapping from ASL to English.

A case in point is adverbs of manner. In written or spoken English, the perception of a verb can vary due to the presence of lexical items such as an adverb of manner (Friedman 1975). Such an adverb may be added after the main verb or after the object in order to express how the action takes place (Valli and Lucas 2000).

Below are two examples of adverbs of manner used in English. Each example consists of two sentences: the first sentence has no adverb of manner and the second contains an adverb of manner. In the first example, the adverb *quickly* describes the manner in which the boy ran. The second example *beautifully* describes the manner in which the couple danced.

- The boy ran.
- The boy ran quickly.
- The couple danced.
- The couple danced beautifully.

In contrast to English, adverbs of manner in ASL are not necessarily expressed by the addition of an independent lexical sign. Adverbs of manner occur by modifying the "quality of the motion" of the verb and are considered nonmanual (Thomson and Martinet 1980, Baker and Cokely 1991, Kluwin 1981, Schnepp et al. 2012, Johnston and De Beuzeville 2016, Padden 2016). These modifications to a lexical sign are used as a way to provide more information. For instance, in English, words such as stomp, tiptoe, slowly, or gracefully may be used to describe a specific manner in which someone walks. In ASL, the sign WALK can be modified to indicate these different manners through intensity and affect (Fischer and Gough 1978).

The addition of other co-occurring nonmanuals, such as facial expressions, may also be used to provide further information in the context of manner (Bahan 1997; Braem and Brentari 2001; Emmorey 2001; Quinto-Pozos 2011).

3 Related work

3.1 Gesture motion

Previous research on the use of adverbs of manner in ASL is limited. Therefore, the starting point for related work is grounded in the analysis of gesture motion. Two important examples of previous work in the area of quantifying gesture motion are the EMOTE (Expressive MOTion Engine) (Chi et al. 2000; Zhao et al. 2000a, b) model and GRETA (Hartmann et al. 2005). EMOTE stems from the Laban Movement Analysis (LMA) and the Effort-Shape model, which draws on LMA's classification of motion in two ways: Effort, which provides qualitative descriptions of energy in motion, and Shape, which describes how the body moves. Table 1 provides qualitative examples of the four parameterizations of Effort proposed by EMOTE: Space, Weight, Time, and Flow.

GRETA's expressivity parameterization expands on EMOTE's techniques for synthesis with six attributes: Overall Activation, Spatial Extent, Temporal Extent, Fluidity, Power, and Repetition.

Although researchers have examined the effects of affect on gesture (Kleinsmith and Bianchi-Berthouze 2012) and Zhao suggests that the EMOTE system would be useful in synthesizing sign languages (Zhao et al. 2000a, b), no one has reported on using such approaches for portraying signed adverbial modifiers.

Sign language synthesis requires more specificity than what is outlined in either EMOTE or GRETA. The characterizations proposed by EMOTE do not fully capture the adverbial modifiers used in ASL. For example, the EMOTE model starts with a base animation that is generic and expressionless in motion, then applies various animation techniques to establish motion parameters such as including arm trajectory, timing, and flourishes for expressive movement. Based on the Effort element, a method of animation would be chosen. For example, EMOTE would characterize this adverbial modification:

$$\frac{\textit{slowly}}{\textit{WALK}}$$

as Bound or Sustained. These attributes would alter the neutral animation by extending the number of keyframes between poses and applying an element of overshoot. As discussed in the following sections, data collected in this study demonstrates that the signed motion for *slowly* WALK does not conform to this EMOTE descriptor. Because these systems are not specialized for sign languages, they do not consider important factors specific to such languages including the signing space and positioning of hands.

A more complete motion model is necessary to allow a 3D avatar to modify signs such as verbs while supporting and respecting ASL's grammatical structure. Accuracy and naturalness in the generated motion are necessary to make sentences as easy to understand as possible. There is also a need for animation techniques to build realism.

3.2 Nonmanual signals

Although manual methods for portraying signs such as hand shape, hand position and palm orientation account for most of the articulated content, grammatical completeness also requires the inclusion of other nonmanual signals. These may include the motion of the head and upper torso, as well as changes to facial expression. Within ASL, these nonmanual signals function on prosodic, lexical, and

Table 1 Effort elements as defined by EMOTE (Chi et al. 2000)

Effort element	Opposing attributes	Descriptions
Space	Indirect	Flexible, meandering, wandering, multi-focus
Attention to the surroundings	Direct	Single focus, channeled, undeviating
Weight	Light	Buoyant, delicate, easily overcoming gravity, marked by decreasing pressure
Attitude towards the impact of one's movement	Strong	Powerful, having an impact, increasing pressure into the movement
Time	Sustained	Lingering, leisurely, indulging in time
Lack or sense of urgency	Sudden	Hurried, urgent
Flow	Free	Uncontrolled, abandoned, unable to stop in the course of the movement
Amount of control and bodily tension	Bound	Controlled, restrained

syntactic levels (Liddell 1978; Baker-Shenk and Padden 1978). For example, the face is required for both emotional and grammatical contexts. One of these is the position of the eyebrows differentiating between wh- questions (who, what, when where, or why), yes/no questions, and rhetorical questions (Baker-Shenk and Padden 1978). In the context of this research, adverbial constructs for adverbs displaying affect are conveyed by emotional facial expressions in accordance with the universal facial expressions laid out by Ekman and Friesen (1971). These adverbial facial expressions showing emotion can be co-occurring with the signing and other nonmanual signals (Weast 2008).

Current avatar animation systems allow for the synthesis of the co-occurring emotion and nonmanual facial signals. In particular, layering wh-type nonmanual signals with emotional facial expressions has had successful results in efforts to study perceived affect (Schnepp et al. 2012). However, this work did not address the use of nonmanual signals in portraying adverbs of manner.

For a successful English to ASL synthesizer, all components of the language should be incorporated. While the previously mentioned studies have been useful for their research of human motion and perceived emotion, they do not account for the application within avatar systems or the need to portray adverbs of manner.

4 Avatar technology

Portraying adverbs of manner on an avatar requires animation to be created for rendering. It also requires that an avatar has finely articulated hands for the accurate generation of handshapes, expressive arms and body to achieve naturalness in motion, and controllable facial expressions for nonmanual signals. Adverbs of manner are portrayed through motion. There are three approaches to apply this motion through animation: manual keyframe animation, motion capture, and procedural animation.

Manual keyframe animation requires an animator to set keyframes showing significant key poses within an action. This is done by adjusting joint rotations and positions of the avatar. This produces expressive, natural results through the animator's ability to exaggerate the motion. For manual animation, keyframe data requires manual intervention by the animator, making it difficult to modify the animation to communicate something new. It can be time-consuming and cost-prohibitive to change the information on all keyframes manually.

Motion capture involves placing sensors at key joints on a signer and recording the sensor's position with a camera at a high sampling rate. This results in a collection of recorded points for each joint in 3D space tracking the signer's motion. This data can be applied to animate an avatar, allowing it to move as the signer moved. Motion capture reproduces the subtle details creating smooth, natural motion, but does not facilitate easy editing. Modification of the high amount of recorded data is difficult. Maintaining the subtleties of the motion would require modification of every angle at every frame.

Procedural animation uses algorithms to drive the animation of signs (Delorme et al. 2009). Such a system needs to take kinetics into account to generate convincing movement and requires a deep understanding of the human figure and how it moves to create a successful animation (Rose et al. 1998). Procedural animation has the benefit of working in real-time and being scalable, though the ability to produce natural motion remains elusive. Its advantage is its flexibility, allowing for easy editing through the use of sparse keyframes and modification of sentence structure and individual signs.

The work reported in this paper uses a hybrid approach that combines the benefits of the manual keyframe animation and procedural animation methods of synthesis. Geometric and timing data for specific beginning and end poses are stored in sparse keyframe databases allowing for interpolation between poses. Having a sparse keyframe database comprised of set poses allows an artist to generate realistic signs by setting the poses and interpolating the transitions. However, with only the pose information, interpolation lacks the subtle movements that occur during transitions. Procedural animation can add generated movements that become more natural through controls that modify them. There are also algorithmic overlays that can enhance the liveliness and natural movement.

The approach outlined below uses appropriate technology that supports the complex needs of sign language portrayal, creating a hybrid animation model to automate the application of adverbs of manner to signed verbs.

5 Procedure

Developing a model to apply the motion modification and additional nonmanual signals to signs incorporating adverbs of manner required several steps, including the selection of adverbs, a motion study and observation of additional nonmanual signals used, animating and validating the exemplars, and data analysis to determine the needed variables to include.

5.1 Adverb selection

The first step was to select what adverbs of manner would be used for this study. Four commonly used adverbs of manner, *quickly*, *slowly*, *happily*, and *sadly*, were chosen to represent different qualities of adverbial modifications to a sign. These adverbs were also chosen because they have corresponding independent lexical items in the adjective form: QUICK, SLOW, HAPPY, and SAD. The focus is on two pairs of contrast: intensity and affect. The adverbs *quickly* and *slowly* contrast in intensity, and the adverbs *happily* and *sadly* contrast in affect. The sign WALK is an appropriate choice for the base verb of the adverbial modification as it is a two-handed sign with repeating motion and no contact.

5.2 Motion study

The motion study provided examples of how the selected adverbs were used in ASL. Initial searches for video references, including the National Center for Sign

Language and Gesture Resources (NCSLGR) Corpus, incorporating the selected adverbs of manner showed a paucity of data demonstrating the adverbs in question, meaning the study would require additional video recordings. Two recording sessions were conducted with the consultation of an ASL interpreter. This interpreter learned ASL at a young age from Deaf signers, is consistently admired for the “Deafness” of her signing and holds a National Interpreter Certification Master credential.

Prior to the recording, the interpreter was presented with the list of sentences to be recorded in written English and was informed of the purpose of the sessions. During the session, the recordings were taken from the front. The interpreter viewed projections of the sentence on a wall in front of them, as shown in Fig. 1.

Sentences were shown one at a time. A high-definition camera recorded the video at 60 frames per second (fps) at 720p. Multiple recordings of each sentence were taken to allow for a more natural performance. The interpreter was given a chance to review the recordings and re-record as necessary. Two sentence variations were recorded to include the use of the adverb of manner through modification of WALK and the use of an adjective lexical item.

$$MAN \quad \frac{happily}{WALK}$$

MAN HAPPY WALK

This provided a starting place for data analysis and prototyping the model. These video recordings were used as the basis for the characterization of motion modifications and additional nonmanual signals. Video recordings were then processed using motion tracking tools within Adobe After Effects to collect data on the change in position, timing, and velocities of joints.

Animators used the collected data, along with observations gained through the visual study of the videos, to recreate the signed sentences using a 3D avatar in the Sentence Generator software (McDonald et al. 2017). The same ASL interpreter then reviewed these initial animations for grammaticality and naturalness. The interpreter then either signed off on the animation or provided clarifying feedback to improve the animation. Two review sessions resulted in the approval of all the synthetic sentences.

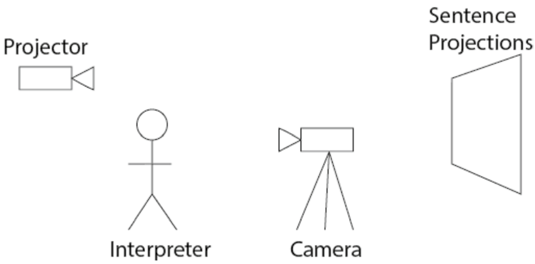


Fig. 1 Recording setup used to capture signed sentenced used as reference

5.3 Data analysis for motion modification

Each animation of WALK with motion modification was created by an animator using keyframe data. Animations were generated by interpolating the in-betweens of the keyframes; this interpolated motion data constructed the basis for the analysis. Data collected directly from the animations resulted in 44 variables including:

- Adverb
- Time in seconds
- x, y, and z position data for:
    - Left and right wrist
    - Left and right elbow
    - Left and right shoulder
- x, y, and z rotation data for:
    - Left and right wrist
    - Left and right elbow
    - Left and right shoulder

This was combined with the following calculated values:

- Displacement for the:
    - Wrist
    - Elbow
    - Shoulder
- Change in velocity for the:
    - Wrist
    - Elbow
    - Shoulder

This motion path data was collected at a rate of 60 frames per second.

With the 44 data variables listed above, the next step was to consider dimensionality reduction. Linear Discriminant Analysis (LDA) (Eisenbeis and Avery 1972) is a commonly used technique for this purpose. An LDA analysis applied to the motion path data resulted in identifying the primary contributing variables. The LDA showed a high degree of separation of the adverbs indicating that there was a certain combination of select variables that could correspond to each adverb of manner, with the first linear discriminate accounting for over 98% of the separation. The primary variables that accounted for the differences in the adverbs of manner included the wrist position and velocity. Figure 2 shows clear

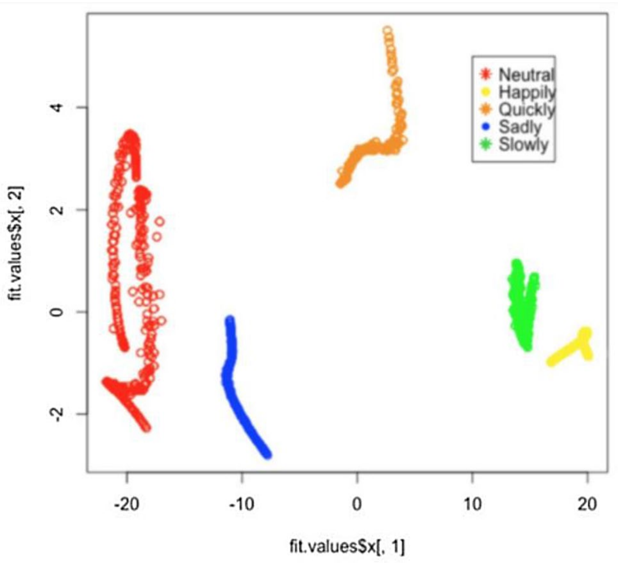


Fig. 2 Separation of adverbs through the first two discriminate functions

separation using the first two discriminant functions for the data gathered from the four adverbs and a neutral sign.

Figures 3 and 4 show the distinct motion paths for the right wrist in the five animations on the transverse plane (x, y), as looking down at the signing space, and sagittal plane (y, z), as looking at the side of the signing space. The color variation through the motion path accounts for the change in velocity.

Fig. 3 Motion path of the right wrist in the traverse plane

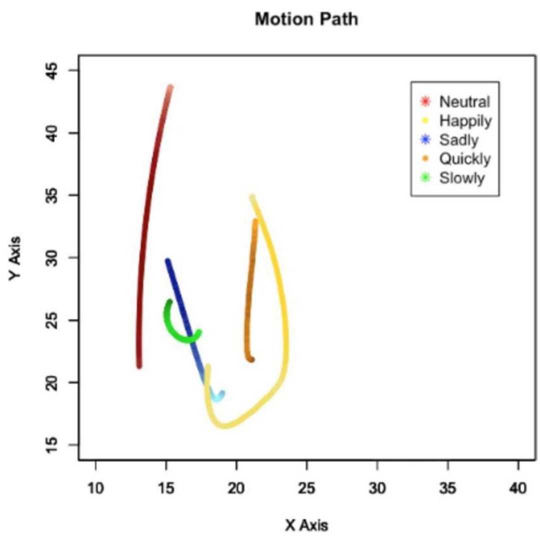
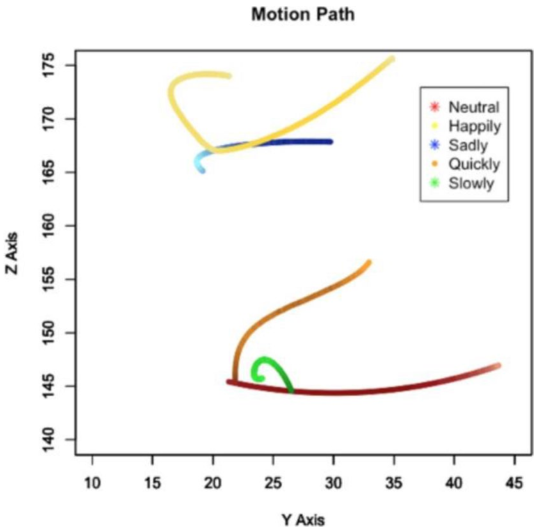


Fig. 4 Motion path of the right wrist in the sagittal plane



6 Developing a model for modifications for adverbs of manner

Based on the LDA results, a hybrid animation model for adverbial modification including motion modification and application of nonmanual signals was then applied to the neutral animation of WALK for initial comparison and revision. Coefficients for each adverb were calculated for adjusted timing, joint rotations, and joint positions to procedurally modify the neutral animation's keyframes. Coefficients for timing accounted for the differences in speed along the motion path shown. The adjustments to timing would compress or lengthen the amount of time between keyframes based on the data collected from the motion study. Arm joint positions were adjusted, changing the overall motion paths for the wrists. Wrist rotations were also adjusted, allowing for fingertips to more naturally follow along the arc of the motion path.

For *slowly* and *quickly*, the data analysis showed that the highest contributing factor for separation was the time variable. Wrist rotations were another contributing part of the motion model for these two adverbs of manner. For both, the direction of wrist rotation tended to follow the rotation of the elbow. In the WALK analysis, *quickly* showed radial and ulnar deviation as the forearm from the elbow moved out and back, whereas *slowly* showed extension and flexion rotation as the arm moved up and down. The LDA analysis also showed that for *quickly*, the wrist rotation would be more significant to the movement than elbow rotation or wrist position.

Though wrist position and velocity were the significant contributors, those alone were not enough to convincingly convey the chosen adverbs. An examination of the *slowly* modifier demonstrates a need to include additional variables and animation techniques. When only the wrist position and velocity were implemented, the sign appeared Bound as described by the EMOTE parametric characterization, but based on the data and the visual observations, *slowly* required expressive joint rotations that could be described as Light. The model was further revised to incorporate animation techniques such as ease-in and ease-out for joint positions.

Ease-in and ease-out are both uneven incremental progressions of motion used to increase the perceived naturalism of an animation (Johnston and Thomas 1981). Ease-in produces an animation that starts *slowly* and accelerates, whereas ease-out is fast at the beginning but decelerates at the end. Consider starting and stopping a car while driving. It will take a little while to accelerate and reach a steady speed, easing into it. Likewise, when the driver applies the brake, the car will decelerate before coming to a stop, easing out from a steady speed. Adding these allowed for changes to the speed of the animation (Burtnyk and Wein 1976).

The changes to the set of contrasting adverbs of manner *happily* and *sadly* relied heavily on a change in the motion path to convey the modification and even more on nonmanual signals making them more complex than quickly and slowly. For both *happily* and *sadly*, the timings were extended along with the use of a slow ease-in and a slow ease-out. The significant difference between the two came from the changes to the motion path. For *happily*, the model lifted and expanded the motion path. For *sadly*, the model lowered and compressed the motion path.

Wrist rotations showed to be a contributing part of the model for both *happily* and *sadly*. For *happily*, the trend of rotation followed what was observed with *quickly* and *slowly* but was more pronounced and extended through the fingers. Whereas, for *sadly*, the wrist rotation was more muted and did not extend significantly in a positive direction.

The model was further extended to include the spine's proximal joints to alleviate the unnatural, exaggerated extension of the arms as the path of the wrist moved further from the body. This aligns with the migration of motion between distal and proximal joints as described by (Brentari 1998).

Additional nonmanual signals were also included as layers to co-occur with the motion modification of the sign. For the extent of this prototype model, nonmanual signals include changes in the face and additional changes to the motion of the spine and head. These nonmaunal signals were especially prominent for adverbs of manner representing affect.

The adjustments provided a hybrid animation model for adverbs of manner that was satisfactory to conduct a user study and apply to other animated verbs.

7 Evaluation design of adverbial modifications to WALK

This section introduces the evaluation design of the adverbial modification model applied to WALK through a user survey. There are two goals of this user study. The first is to determine which portrayal of adverbs of manner are judged to be the most understandable:

- Lexical item only
- A combination of lexical item and motion modification
- Motion modification only
- A combination of motion modification and additional nonmanual signals

- A combination of lexical item, additional nonmanual signals, and motion modification

The second goal is to determine whether using multiple channels will increase the intensity of the perceived adverb. In other words, is the adverb *sadly* portrayed by motion modification and additional nonmanual signals perceived as sadder than when the adverb is portrayed by motion modification alone? The video (0:36) at the following link shows the animations used in the survey for the adverbial modification *sadly*: http://sltat.cs.depaul.edu/2019/Moncrief.mp4.

This study uses four common adverbs of manner, *quickly*, *slowly*, *happily*, and *sadly*, to determine the effectiveness of these modifications. Users view four animation sets and judge them.

The evaluation participants were individuals over the age of 18 and proficient in ASL. Qualtrics hosted the evaluation online in a survey format with the informed consent and instructions accessible in ASL. Animations were embedded into the survey for the participant to view. The format allows for re-watching of videos as many times as the viewer chooses.

Participants were shown randomized sets of animations comparing different treatments to an adverb as a side-by-side A/B test. For each comparison, participants responded to four questions:

- Which animation is the *saddest/happiest/quickest/slowest*?
- Which animation is the most natural?
- Which animation is most grammatically correct?
- Which animation shows the best head and body coordination?

Before launch, the survey was reviewed by a certified ASL interpreter and the Institutional Review Board at DePaul University (IRB# RM020617CDM).

8 Interim user study results

Early results indicate that all respondents showed a strong preference for animations that displayed nonmanual signals for adverbs involving affect (*happily*, *sadly*). For *happily*, the animation that was perceived as the happiest used a combination of all three channels: the lexical item, the adverbial modification, and the nonmanual signals. This animation was also perceived as the most natural, grammatically correct, and having the best head and body coordination. In contrast, for *sadly*, the preference for all four questions was the animation with the adverbial modification and nonmanual signals but excluding the lexical adjective.

For *quickly*, respondents showed a preference for the animation containing the adverbial modification and nonmanual signals but excluding the lexical item. For the adverb *slowly*, a clear preference was not evident. The two most preferred animation choices included all three channels and the animation containing the lexical adjective and the adverbial modification. For all four adverbs, there was little to no preference for animations that displayed just the lexical adjective or just the adverbial modifier without nonmanual signals.

Based on the current results from the user study, the hybrid animation model has been consistently applied to other signed verbs using both the motion modification and nonmanual signals.

9 Generalizing the model

A variety of other verbs were chosen to test the model's limitations and evaluate the generalizability of this prototype. To this end, the following sign attributes were considered: one-handed or two-handed, varying use of contact or no contact, and repeating or non-repeating motion. The selections are shown below in Table 2 including the sign used for model development, WALK, for comparison.

It was also important to evaluate signs that displayed a different motion path than that used for the model to ensure that the motion modification was adaptable and maintained naturalness in appearance. A neutral version of WALK alternates the extension and retraction of the arm, led by the hand, in a flat motion path that does not vary up or down. The signs chosen for further consideration of the model were again chosen for their variety of motion paths compared to WALK. This includes variations in circular motion, high to low motion, and lateral motion.

9.1 One-handed vs. two-handed signs

Some signs in ASL are formed with one hand, while others are formed with both. For example, WALK uses both hands alternating extending in and out, where GIVE uses the signer's dominant hand. The model needed to recognize if a sign was one-handed or two-handed. When concatenating signs to string a sentence together, for instance, a one-handed sign followed by a two-handed sign, the model could impact the interpolation from one sign to the next. This could cause the hands to move in a way that was not part of the original sign and should not be a part of the modification. The model now recognizes if the sign is one-handed or two-handed based on a comparison between keyframes throughout the sign. This comparison looks at the

Table 2 Chosen signs and their characteristics for selection

Sign	Two-handed	Contact	Repeating
WALK	X		X
ASK			
GIVE		X	
THINK			X
BREATHE	X	X	X
INFORM	X		
CLEAN	X	X	
SEARCH	X		X

x, y, and z position information for the shoulder, elbow and wrist. If there is not a change from one keyframe to the next, no modification is applied.

9.2 Non-repeating vs. repeating motion signs

Non-repeating motion signs are signs that do not repeat all or a section of the motion path, such as GIVE that starts at the chest and moves forward and away from the body. In contrast, repeating motion signs are signs that repeat all or part of the motion path, such as WALK. When applying the model to non-repeating motion signs, there is less overall movement to aid in the perception of the adverbial modification. The adjustment of surrounding signs can help with the perception of the adverbial modifiers of intensity, as mentioned in Sect. 9.4 with *quickly*. In the case of affect, the use of nonmanual signals (facial expressions) play an important role.

9.3 Contact signs

Contact signs are signs that come into contact with the signer and may include contact between hands, the face, or other parts of the body. For example, the sign BREATHE has both hands come in contact with the signer's chest. The model did not initially accommodate signs that contact the body, such as BREATHE or CLEAN. In several cases, this resulted in the hands overshooting and ending in a collision with the body at the point of contact. To account for this, the neutral animations needed to be adjusted to include a tag on the keyframe with the contact. With this tag in place, the model would negate any positional/rotational motion path modifications applied to that keyframe. Adjustments would then only impact the surrounding keyframes, leaving the contact as originally animated.

In cases of large inward motions where the original sign does not have contact with the body, the model needs to account for potential collisions that the modifications may cause. For instance, since *happily* exaggerates the sign, there may be rare cases that cause the hand to penetrate the avatar's chest or head. The modified wrist position is compared and reduced when needed.

9.4 Adverbs of manner—intensity

The adverbial modification for *slowly* proved to be the easiest to transfer to the chosen signs shown in Table 2, followed by *quickly*. Unlike *happily* and *sadly*, the changes to this set of contrasting adverbs of manner relied less on a change in motion path to convey the modification and even less on nonmanual signals such as facial expressions. Whereas the model for *slowly* was obvious across all evaluated signs due to the duration of the sign being extended, *quickly* was noticeable only on signs that had repeating motion, such as BREATHE. The motion modification was not obvious for nonrepeating signs such as ASK and GIVE likely due to the relatively short duration of the neutral version of nonrepeating signs.

Upon application of the hybrid animation model on additional verbs to modify for *quickly*, the overall timing of the sign was not changed significantly. Further

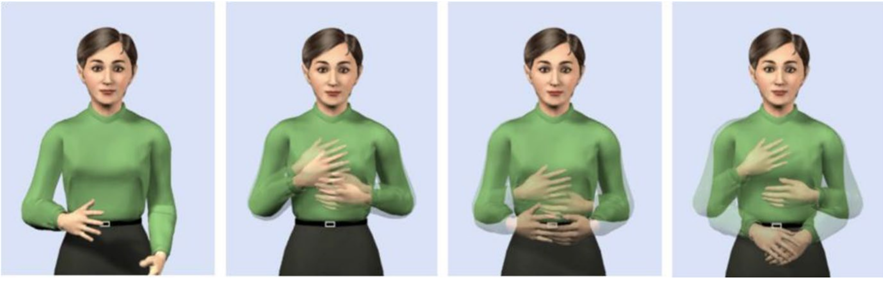


Fig. 5 An overlay with the neutral animation of BREATHE and BREATHE *slowly*. Follow this link to watch the full animation (0:03): http://sltat.cs.depaul.edu/2019/Moncrief.mp4

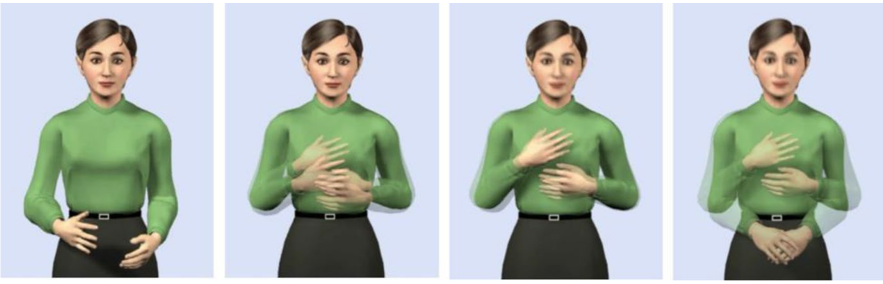


Fig. 6 An overlay with the neutral animation of BREATHE and BREATHE *quickly*. Follow this link to watch the full animation (0:18): http://sltat.cs.depaul.edu/2019/Moncrief.mp4

increasing the speed only made the animation come across as less natural. The motion modification is applied to the surrounding signs to account for this. In the case where the adverbial modifier for *quickly* was applied, the prior and following signs were also modified in the same way. Figures 5 and 6 visualize the differences in duration overlaying a neutral animation with one having motion modification applied. Figure 5 shows an overlay of the neutral animation for BREATHE and the applied *slowly* model. Figure 6 below shows the comparison with *quickly*.

As this model is applied to other verbs, the importance of the wrist rotation for *quickly*, variers based on the path of the wrist. As mentioned on Sect. 6, the LDA analysis showed that for *quickly*, the wrist rotation was found to be significant. For instance, GIVE *quickly*, a short one-handed nonrepeating sign, may see more extension of the wrist rotation than the elbow. This will be explored by incorporating more verbs as a part of this study.

9.5 Adverbs of manner—affect

As mentioned in Sect. 6, modification for happily and sadly relied on changes to the motion path and the inclusion of nonmanual signals. Figure 7 shows an overlay of the neutral GIVE and GIVE with the *happily* modification. This demonstrates expanding of the motion path for *happily*. For *sadly*, the motion

Fig. 7 An overlay with the neutral animation of GIVE and GIVE *happily*. Follow this link to watch the full animation (0:29): http://sltat.cs.depaul.edu/2019/Moncrief.mp4

modification contrasts *happily* by lowering and compressing the motion space. Based on the initial data collection, *sadly* showed to have a continuous lowering effect on the signing space in signs with repeating motion. Figure 8 demonstrates this. Keyframe data were compared in the neutral animation to see if the sign comes back near the starting position. Any keys after would have an increased drop in their position and an even slower timing adjustment.

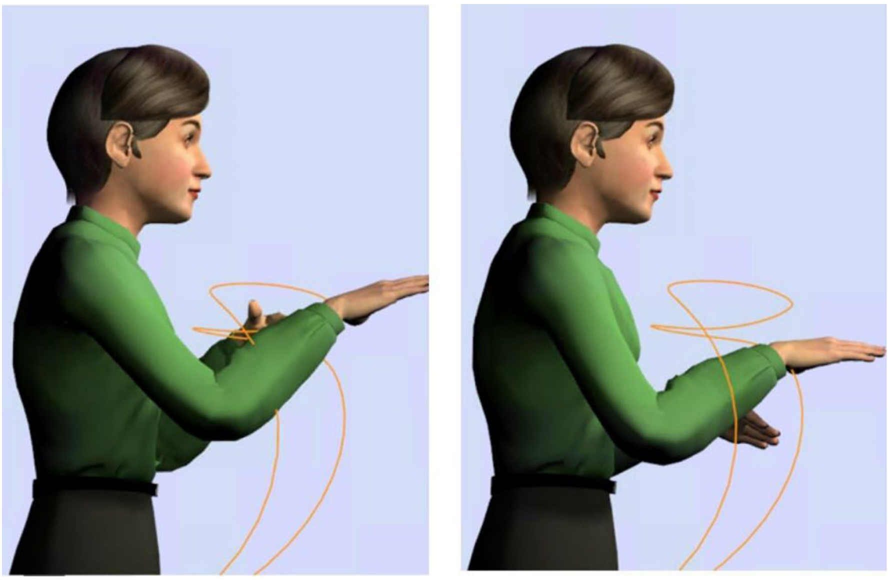


Fig. 8 The motion path of the right wrist for WALK *sadly*. The first extension is shown higher, with repeating extensions shown lower

10 Conclusion and future work

Creating a generalized model for the portrayal of adverbs of manner on an avatar is a complex undertaking requiring a unique interplay of anatomy, motion, and expression. Interim results from the user study are showing a need for the inclusion of additional nonmanual signals, such as facial expressions, and that there is a preference for motion modification applied to lexical verbs as opposed to sentences with the lexical adjective and neutral verb. These results may not seem surprising as these modifications create more expressive animations, but they also confirm that the procedurally generated motion modifications applied to the verb WALK are acceptable to the users. These results are promising and show merit in continuing the development of procedural techniques to include adverbs of manner through motion modification and application of nonmanual signals.

Generalizing the model of adverbial modification required consideration of several aspects of the sign, including motion, nonmanual signals, and sign attributes. These extensions are not considered a complete solution but did provide a method to examine the limits of migrating modeled adverbial modifications from one sign to another.

This model will be strengthened by incorporating observations from additional lexical signs with adverbial modifications applied to them. This will allow for a more accurate expansion of the model to other signs. The addition of other adverbs of manner will be examined as well. Exploration of various data sources and tools such as OpenPose, will also be considered for improving the model. This will include the analysis of mocap data and further user testing.

References

Bahan BJ (1997) Non-manual realization of agreement in American Sign Language

Baker-Shenk CL, Cokely D (1991) American Sign Language: a teacher's resource text on grammar and culture. Gallaudet University Press, Washington, DC

Baker-Shenk CL, Padden C (1978) Focusing on the nonmanual components of ASL. In: Siple P (ed) Understanding lanugage through sign language research: perspectives in neuroloinguistics and psycholinguistics. Academic Press, New York, pp 27–57

Barraquand J, Latombe J-C (1991) Robot motion planning: a distributed representation approach. Int J Robot Res 10:628–649

Braem PB, Brentari D (2001) Functions of the mouthing component in the signing of deaf early and late learners of Swiss German Sign Language. In: Brentari D (ed) Foreign vocabulary in sign languages: a cross-linguistic investigation of word formation. Lawrence Erlbaum, Mahwah, pp 1–47

Brentari D (1998) A prosodic model of sign language phonology. MIT Press, Cambridge

Burtnyk N, Wein M (1976) Interactive skeleton techniques for enhancing motion dynamics in key frame animation. Commun ACM 19:564–569

Chi D, Costa M, Zhao L, Badler N (2000) The EMOTE model for effort and shape. In: Proceedings of the 27th annual conference on computer graphics and interactive techniques.

Delorme M, Filhol M, Braffort A (2009) Animation generation process for sign language synthesis. In: Conference on advances in computer-human interactions, IEEE, pp 386–390

Eisenbeis RA, Avery RB (1972) Discriminant analysis and classification procedures: theory and applications. DC Heath, Lexington

Ekman P, Friesen WV (1971) Constants across cultures in the face and emotion. J Pers Soc Psychol 17(2):124
Emmorey K (2001) Language, cognition, and the brain: insights from sign language research. Psychology Press, Hove
Fischer S, Gough B (1978) Verbs in American sign language. Sign Lang Stud 18:17–48
Friedman LA (1975) Space, time, and person reference in American Sign Language. Language 51:940–961
Hartmann B, Mancini M, Pelachaud C (2005) Implementing expressive gesture synthesis for embodied conversational agents. In: International Gesture Workshop
Johnston T, De Beuzeville L (2016) Auslan corpus annotation guidelines. Auslan Corpus
Johnston O, Thomas F (1981) The illusion of life: Disney animation. Disney Editions, New York
Kleinsmith A, Bianchi-Berthouze N (2012) Affective body expression perception and recognition: a survey. IEEE Trans Affect Comput 4:15–33
Kluwin TN (1981) A rationale for modifying classroom signing systems. Sign Lang Stud 31:179–187
Liddell S (1978) Nonmanual signals and relative clauses in American Sign Language. In: Siple P (ed) Understanding language through sign language research. Academic Press, New York, pp 55–90
Malala VD, Prigent E, Braffort A, Berret B (2018) Which picture? A methodology for the evaluation of sign language animation understandability. In: Multimodal signals: cognitive and algorithmic issues. pp 83–93
McDonald J, Wolfe R, Alkoby K, Carter R, Davidson MJ, Furst J, Hinkle D, Knoll B, Lancaster G, Smallwood L, et al (2005) Achieving consistency in an FK/IK interface for a seven degree of freedom kinematic chain
McDonald J, Wolfe R, Wilbur RB, Moncrief R, Malaia E, Fujimoto S, Baowidan S, Stec J (2016) A new tool to facilitate prosodic analysis of motion capture data and a data-driven technique for the improvement of avatar motion. In: Proceedings of language resources and evaluation conference (LREC)
McDonald J, Wolfe R, Johnson S, Baowidan S, Moncrief R, Guo N (2017) An improved framework for layering linguistic processes in sign language generation: why there should never be a "brows" tier. In: International conference on universal access in human-computer interaction, pp 41–54
Padden CA (2016) Interaction of morphology and syntax in American Sign Language. Routledge, London
Quinto-Pozos D (2011) Teaching American Sign Language to hearing adult learners. Annu Rev Appl Linguist 31:137
Rose C, Cohen MF, Bodenheimer B (1998) Verbs and adverbs: multidimensional motion interpolation. IEEE Comput Graph Appl 18(5):32–40
Schnepp JC, Wolfe RJ, McDonald JC, Toro JA (2012) Combining emotion and facial nonmanual signals in synthesized american sign language. In: Proceedings of the 14th international ACM SIGACCESS conference on computers and accessibility
Stokoe W (1960) Sign language structure. Studies in Linguistics, occasional papers 8. Buffalo, New York
Thomson AJ, Martinet AV (1980) A practical English grammar. Oxford University Press, Oxford
Valli C, Lucas C (2000) Linguistics of American sign language: an introduction. Gallaudet University Press, Washington, DC
Weast T (2008) Questions in American Sign Language: a quantitative analysis of raised and lowered eyebrows. Doctoral dissertation, The University of Texas, Arlington
Zhao L, Costa M, Badler NL (2000) Interpreting movement manner. In: Proceedings of the computer animation 2000, pp 98–103
Zhao L, Kipper K, Schuler W, Vogler C, Badler N, Palmer M (2000) A machine translation system from English to American Sign Language. In: Conference of the Association for Machine Translation in the Americas