

An Improved Framework for Layering Linguistic Processes in Sign Language Generation: Why There Should Never Be a “Brows” Tier

John McDonald^(✉), Rosalee Wolfe, Sarah Johnson, Souad Baowidan, Robyn Moncrief, and Ningshan Guo

DePaul University, Chicago, IL, USA

{jmcDonald,wolfe}@cs.depaul.edu, sarahej101@gmail.com,
{sbaowida,rkelly5,nguo4}@mail.depaul.edu

Abstract. Creating legible animations of sign language that satisfy the needs of native users requires drawing from the fields of computational linguistics, computer animation and user experience research. This paper explores the problem of layering the many linguistic processes that contribute to a sign language utterance in avatar animation. A new framework is presented that satisfies requirements from each of these fields, and yields a flexible architecture for sign language avatars capable of leveraging a wide range of animation techniques to generate rich multi-layered animations of sign.

Keywords: Sign language avatars · Computational linguistics · Computer animation · Hierarchical models · User interface

1 Introduction

In many countries access to spoken and written language remains an extreme challenge for those who use sign languages as their preferred means of communication. Automatic translation technologies have the potential to help bridge this gap. Because sign languages are expressed on the human body and face rather than in a written form, avatars capable of displaying the full naturalistic motion of sign language are essential for such translation systems.

In this paper, the term *sign language* will be understood to refer to all fully-qualified linguistic systems (independent languages) that are visual/gestural in their modality. This distinguishes these languages from other signed communication systems such as Signed English [1]. Unlike spoken or written languages, sign languages communicate in multiple channels concurrently. The hands and arms often form the primary manual channel that conveys lexical information, but additional signals co-occur on the torso, neck and face that can add linguistic and paralinguistic information to intensify, invert or otherwise modify the meaning of this information [2].

Current sign language avatars are capable of displaying a stream of lexical units [3, 4], but how best to represent and coordinate information beyond the manual channel has proved to be a challenge, and is still an open question. This paper presents a novel

framework for layering linguistic processes for avatar technology that facilitates greater expressivity, naturalness and legibility beyond the manual channel. This new framework allows for both synchronous and asynchronous coordination of processes among channels while avoiding the robotic motion often associated with avatars. In addition, the framework provides an elegant and parsimonious method to facilitate a high degree of flexibility in the way each channel can influence the avatar's motion.

Developing this framework required a multidisciplinary approach. Automated sign language generation draws on a number of disciplines, including computational linguistics, computer animation and user experience which all inform the software engineering design. To motivate the new framework, this paper analyzes sign language avatar technology from each of these perspectives. In the following three sections, we discuss relevant features of each discipline that will impose constraints on the system, which will be addressed in the final section of the paper that lays out the architecture of the system.

These disciplines require a high level of flexibility and naturalness in an avatar framework. In particular, the framework cannot limit a physical feature such as brows to a single linguistic or animation process, yet each process will affect a wide range of physical features on the avatar. Additionally, the avatar framework cannot be limited to a single method of animation. The implementation presented is an elegant solution to enable this flexibility, and it requires little additional code beyond the traditional services provided by animation systems to support key frame, procedural and motion capture animation.

2 Computational Linguistics

Avatars have several promising uses in computational linguistics, and the requirements of these applications yield important priorities for the structure and quality of avatar animation technology. Since an avatar is the most appropriate target for any spoken-to-sign translation system, it must be capable of expressing all aspects of sign language. To date, avatars have struggled to provide the flexibility, fluidity and legibility desired by native signers.

Second, avatars can serve as a helpful component for improving the annotation of video. Applications such as Elan, iLex, and Anvil allow researchers to annotate videos of sign language in multi-tier organizations [5–7], and validation is a continual concern. If annotations are used to drive an avatar automatically, the resulting output can be compared with the initial video for discrepancies. This provides an independent method of verification [8].

Finally, avatar technology shows potential as a hypothesis-testing tool. One possible method is to apply the hypothesis to an avatar, generate animations and then review the animations with the Deaf community. In this capacity, avatars have a distinct advantage over video technologies, since they can allow for a single process to be included or excluded in isolation. It is much more difficult to ask a signer, “please do the exact same thing, but without the Y/N question indication” [9]. However, for avatars to be used in this way, they must produce a more realistic depiction of sign than currently possible.

To create an avatar capable of serving these applications, one must consider the linguistic structures that the avatar must express. As independent, natural languages,

sign languages have grammars that do not correspond directly to the linguistics of spoken languages. Animating sign languages flexibly and efficiently requires drawing on a deep understanding of sign language structure. The discipline of sign language linguistics yields a wealth of information that can be used for this purpose. Linguistics also poses several challenges in the coordination of channels, which heretofore has not been satisfactorily addressed in avatar technology.

Initial findings in the field [10] have been extraordinarily helpful for structuring avatar motion by categorizing the manual parameters of an utterance into handshapes, positions, palm orientation and motions. However, such elemental parameters are only the beginning of the structure that a sign language avatar will need to express. More recent research has revealed rich nonmanual signals on the face and torso [11], a system of interacting referential frames for communicating reported speech [12], and many other structures such as classifier predicates [13]. All of these processes can co-occur, and can even individually involve multiple co-occurring movements.

Further research has also indicated that language processes do not lay claim over a particular subset of human anatomy to the exclusion of co-occurring processes. Instead, a complex structure of linguistic forms will layer onto multiple parts of the body to express the intended utterance [14]. This extends and reinforces initial studies on features of the body which indicate that multiple linguistic processes can combine to influence a single part of the body. Consider the following examples of co-occurring processes.

- Sign languages often communicate syntactic information, such as whether an utterance is a statement or a question, by raising and lowering the brows. In addition, brows can be used to communicate emotional content in utterances such as joy and anger.
- The neck can be used to communicate syntactic information in some sign languages such as the use of a head-nod to enumerate a list in *Langue des Signes Française*. In addition, the neck can be turned when reporting the discourse of third parties.

Such processes will not occur in a synchronous manner. The onset of information in a nonmanual channel may or may not coincide with the onset of lexical items, and the duration of the nonmanual signal may differ from the duration of the lexical item [15]. Thus, multiple linguistic processes can simultaneously influence the position and/or orientation of an individual anatomic feature.

For example, consider a fan of the Cleveland baseball team who asks sadly, “CUBS WIN?” This phrase consists of two lexical signs, along with two co-occurring nonmanual signals. The final blink is a prosodic indicator for the end of the phrase, as presented in (1).



To analyze the layering of linguistic processes in this phrase, we visualize them as a block diagram in Fig. 1.

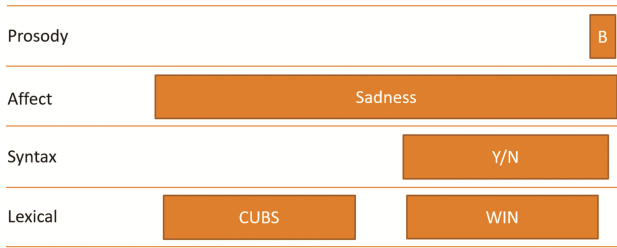


Fig. 1. Layered structure of the utterance “CUBS WIN?”

The processes are manifested as movements or positions on the signer’s body. Multiple processes can affect a single joint or position. Though, the movements or positions are sometimes in conflict, the utterances are legible to a fluent signer. For example, research has shown that multiple processes influence the timing and movement of the brows [16, 17]. Both the question marker and the sadness expressed in (1) are communicated by the brows. In the former, the brows are raised to indicate a y/n question, and in the latter the brows are lowered, communicating sadness. However, the timing and intensity of the two brow transitions allow a signer who views the utterance to identify both processes easily.

For all of these reasons it is clear that a complete framework for communicating sign language should not have a single “brows” tier, as a single tier would have difficulty encoding the multiple influences affecting the brows. This asynchronicity presents several challenges for a framework in the coordination of channels. From a linguistics perspective, the framework should focus on language processes, not anatomy. It must allow any combination of linguistic processes in an utterance, and the sign language avatar technology should provide clear labelling of the linguistic processes. In particular, the framework should satisfy the following requirements based on the analysis in this section:

- (L1) Multiple processes can and will affect the same geometry.
- (L2) Processes may start and end asynchronously.
- (L3) Processes may be enabled and disabled at the user’s discretion.

3 Computer Animation

The quality of an avatar’s signing hinges on the quality of the underlying computer animation. Its usefulness in the previously discussed applications is dictated by the flexibility of the graphical architecture and the types of animation techniques that it supports. Interestingly, the requirements for animating a sign language avatar differ from and are, in some ways, more demanding than those for animating film characters and game avatars.

Avatars for sign languages do share some similarities with animated characters from film and computer games, but they also have additional requirements that set them apart. This can be quite surprising at first blush since realistic animated characters are so ubiquitous in today’s film industry. Viewers expect that realism to carry over to signing avatars. Unfortunately, the two are very different since film is not an interactive nor a

generative medium. The motion of the characters in Pixar's *Toy Story* is the same today as it was when the film was released in 1995 [18]. The film, once rendered, is set for all time, and editing can only be done at the greatest of expense.

Compare this to sign languages which, being productive, can express a functionally infinite range of utterances. If an avatar is to be used for sign language generation rather than simple playback of prerecorded sequences, then the system must support the generation of novel utterances at the whim of the user. Such flexibility is closer to what a user expects from a computer game avatar, but such avatars usually only have a limited number of movements such as swinging a baseball bat and sliding into a base. Further these predefined movements can only be combined in predefined ways [19].

Another difference between signing avatars and animated film/game characters is that cartoon animators utilize labor-saving shortcuts such as using simplified hands consisting of only three fingers and a thumb. Such simplifications would be inappropriate for sign languages, since there would be no way to distinguish between the 7 and 8 handshapes of ASL. Even when animated characters are rigged with realistic hands, there is further complexity in generational avatar systems such as automatic collision avoidance which is necessary for such actions as entering and leaving the letter "T" in ASL fingerspelling, see Fig. 2.



Fig. 2. ASL Handshape 'T'

Several distinct approaches exist for animating avatars all of which have been exploited over the years in various sign synthesis systems including:

1. artist-driven keyframe animation [20],
2. motion capture [21],
3. key-frame synthesis from linguistic descriptions [22],
4. procedural techniques [23].

Each of these methods has its own advantages. Artist driven keyframe animation can be highly realistic, and provides a sparsity of data that can facilitate editing and combining animation clips, but its realism depends largely on the skill of the artist.

Motion capture, can produce extremely realistic motion, as long as the body type of the avatar matches the body type of the recorded person. Unfortunately, motion capture also produces a density of data that causes extreme challenges for editing and combining recorded motions. Doing so relies on large libraries of recorded clips that enable searching not only for the nature of the desired motion but also for the motion at the boundary of the clip in order to smoothly combine them in sequence [24]. Research into this is ongoing.

Reconciling the need to drive an avatar from linguistic data with the demands of producing natural human motion is an ongoing challenge. This is especially true given the requirements of the target audience who expect legible flowing sign, and can find it difficult to read stiff robotic motion. Synthesizing sign exclusively using linguistic descriptions results in such robotic motion. However, such an avatar has the flexibility to combine any lexical items that the linguistics encodes. Conversely, natural and realistic avatar motion relies on extensive animator time or motion capture data at the expense of flexibility. Such systems can only express what has been either animated or pre-recorded.

Procedural animation techniques may also be useful for driving a sign language avatar. For example, consider the fact that the joints in the human body do not start and end their movements simultaneously. In a role shift, the head rotates first, followed by the hips and then the upper spine and shoulders [25]. Such subtleties are already baked into motion capture recordings, but they must be handled manually or modeled procedurally in a key-frame animation system. Experienced animators are skilled in incorporating asynchronicity of this kind, but it is time consuming. Procedural techniques can add such effects and shorten the animator's time and expense [15].

Ideally, an avatar framework would have the capacity to incorporate any and all of the four animation techniques, employing the one best suited for any given language process. Regardless of the underlying animation technique, creating natural, convincing animation requires an acute attention to detail at the biomechanical level, including subtle changes to the avatar's movement that cause no distinguishable change at the linguistic level but affect the legibility of the generated utterance. The framework should support the tuning of such motions within the confines of the linguistic constraints.

Understanding how these four animation approaches can cooperate in an avatar framework requires a deeper analysis. All animation systems will model the human body as a skeleton of articulated bones arranged in a hierarchy so that rotating bones closer to the root of the hierarchy will also affect child bones, see Fig. 3. In this articulated figure, the upper spine, neck, head, shoulders and arms are all children (descendants) of the waist bone. When the waist moves, they also move.

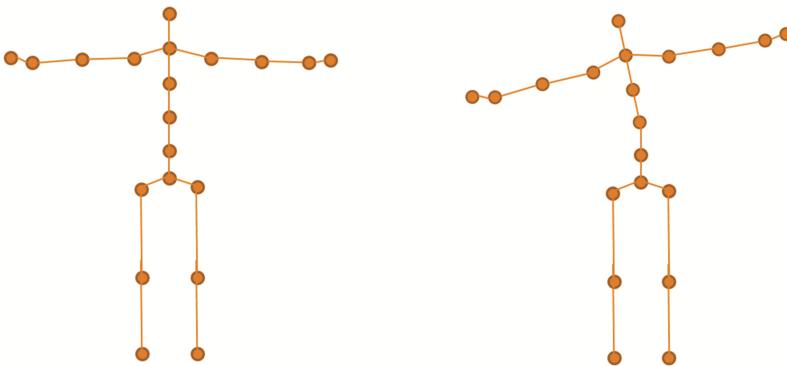


Fig. 3. Hierarchical Skeleton: the waist influences the orientation of all bones descending from the waist.

In the case of a key-frame animation system, a set of controllers for each bone will interpolate a set of key positions or rotations using a variety of methods [26]. The result is then multiplied by the parent’s transformation to get the overall transformation of the bone.

The challenges of using computer animation to produce sign language thus lead us to the following requirements:

- (A1) Multiple processes will affect each bone and combine to produce the final orientation of each bone.
- (A2) Any process that affects multiple bones may have differing start and end times on each bone but will need to be controlled in concert.
- (A3) These processes may require different animation techniques, interpolation schemes or procedural computations.

To better understand how these requirements present challenges for an animation system, consider the sentence in (1). The linguistic processes become layers of animation that must be combined. See Fig. 4. Several of these animation blocks may affect any given part of the body, such as the brows. These include the syntactic marker for a Y/N question and the extralinguistic expression of sadness. Suppose also that in the animation of WIN the artist added a subtle movement of the brows to enhance the legibility and naturalness of the sign, and perhaps also added a colorful nonmanual embellishment associated with the lexical item CUBS.

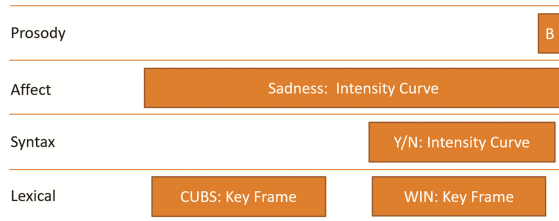


Fig. 4. Animation techniques for different processes in an utterance

On the lexical track we have the two signs CUBS and WIN. These are built as key-frame animations and each consists of a sequence of orientations specified at a subset of the joints. Layered on these lexical items is the Y/N question nonmanual marker which raises the eyebrows. This action begins slightly before the WIN sign and ends slightly after it. The extralinguistic emotion of sadness lowers the eyebrows, combining with the raising from the question. This action encompasses the whole utterance. Finally the eyes blink at the end of the phrase, which may also lower the eyebrows very slightly. Each of these processes is specified as a strength of expression defined by an intensity curve.

Mixing different types of computations on a given bone can be a challenge, as the avatar framework will need to compute the value of each process at each frame and multiply them to obtain the final orientation of the bone. Figure 5 motivates how each of these processes are manifested mathematically. Consider the time t indicated at the end of the production of the sign WIN in the figure. Four processes affect the eyebrows at time t . Each of these processes will create a transformation, M_{procs} affecting the brow as indicated

in the figure. The final transformation of the brow is the combination of all of these effects, some raising and some lowering the brows. Since each is a rotation on a bone that moves the eyebrow, the total transform can be built as a product of these transformations:

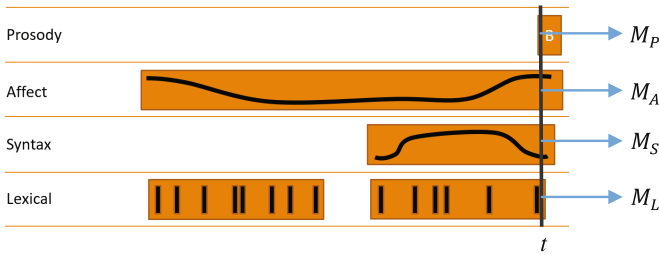


Fig. 5. Transformations that affect the eyebrows

Using the linear algebra convention of pre-multiplication, the final transformation on the brows will be:

$$M_P * M_A * M_S * M_L$$

A system that can manage such processes and combine them properly for a bone will also be able to handle requirement **A3** above since it does not matter whether the underlying representation is procedural, key-frame or motion capture. Each are individually evaluated and combined on the bone. Current sign language avatars are limited by either lacking the capability of layering these processes on some or all of the joints in the avatar's body, or lacking the facility to tune these processes as dictated by linguistic and physical constraints.

The next two sections explore how to manage these processes from the perspectives of user experience and software engineering. In particular, the next section explores how the linguistic and animation requirements will inform the requirements of the user experience.

4 User Experience

The user interface will tie the linguistic descriptions to the animation techniques while also giving necessary control to adjust avatar movement within the linguistic constraints. The interface must accommodate three different types of users:

1. Linguists who will be primarily concerned with the structure of the utterance.
2. Animators who will be primarily concerned with the realism and flow of the avatar's motion.
3. Machine translation researchers, who will want the software to output natural correct sign with as little user intervention as possible.

Animators are concerned with the appearance of an anatomic feature, as contrasted with linguists, who identify the language processes that influence that feature. Returning to the example in (1), consider the difference between raising an eyebrow in an animation

and determining what combination of processes in the language caused the motion. A linguist will want to designate the time of onset and offset of each process. An animator will want to adjust the finer details of timing, including transition shapes (attack and decay) of the envelope as well as adjusting the envelope's steady state. The interface must support the goals of each type of user, while quietly taking care of other details.

The linguistic (**L1–L3**) and animation (**A1–A3**) requirements share commonalities in terms of flexibility for both timing and affected components of the avatar. To satisfy these, the framework's interface must visualize the temporal component as signed utterances unfold over time. Further, this temporal component must be subdivided into the various process tracks or tiers that control the avatar's motion. As in many annotation systems, the new framework displays the time axis horizontally and the tracks organized vertically, see Fig. 6. This organizational scheme is familiar to sign language researchers, and is similar to sign language annotation systems. It allows easy temporal comparison and coordination of elements.

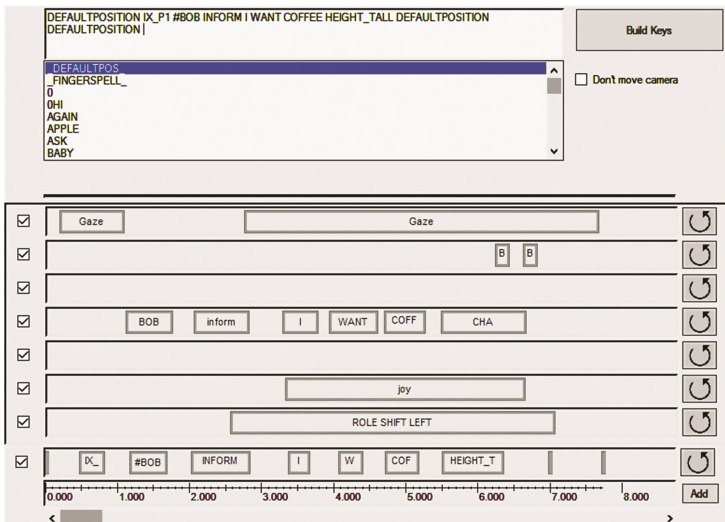


Fig. 6. Paula sentence generator interface

This interface has several features that satisfy requirements **L1-3** and **A1-3** above:

1. A given bone in the hierarchy may be influenced by multiple tracks (**L1, A1**).
2. Timing of animation segments can be controlled independently in each track, and thus tracks may independently control the configuration and timing of multiple sections of the avatar (**L2, A2**).
3. As disparate parts of the human anatomy may be involved in a specific process, bones may not necessarily be contiguous in the hierarchy (**L2, A2**).
4. The check-boxes at the left of each track allow the track to be enabled and disabled at the user's discretion. In addition, tracks may be individually edited without affecting the processes in other tracks (**L3**).

Several manual and nonmanual processes affect the spine of the avatar during the production of LARGE including:

1. The manual channel will raise the shoulders and give a small lean of the spine away from the raised hand. This process uses artist driven key-frame animation.
2. The role-shift used to mark reported dialog will turn the spine along the axis of the body. This process uses a procedure that rotates the body to take the role of a previously indexed discourse participant. This procedure manages the asynchronicity of the spine bones in the shift [15].
3. The emotion of joy will tend to raise the shoulders and arch the spine slightly. This process uses a pre-generated pose controlled by an intensity curve.
4. To increase realism throughout the utterance, a small amount of noise is applied to the joints in the spine to “liven” the avatar. This is a procedural application of Perlin noise [23].

The avatar framework should manage all four of these processes without letting them interfere with each other. Each has its own method of computation at a specific time in the animation. Consider the single bone “Waist” in the hierarchy displayed in Fig. 8. To allow the avatar to combine these four effects, we split this bone into four *sub-bones* as displayed.

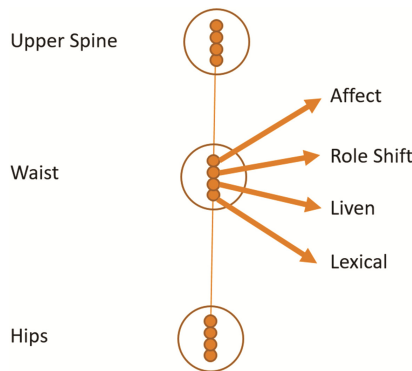


Fig. 8. Bone structure in the torso

Each sub-bone is a fully-qualified bone in the animation engine with its own transformation controller. We call them sub-bones because each has identical positions to the main waist articulator. They are hierarchically organized from parent to child in the order dictated by the tracks in the sentence generator interface seen in Fig. 6. In this case, the hierarchy is:

```

Waist_Lexical
  Waist_Liven
    Waist_RoleShift
      Waist_Affect
  
```

This framework satisfies the linguistic **L1–L3** and animation **A1–A3** requirements in the following ways:

1. **L1** and **A1**: Since each sub-bone has its own animation controller, each will be set and controlled independently. The hierarchical nature of the skeleton automatically combines the effects to produce a final transform on the overall bone.
2. **L2** and **A2**: The timing of effects or processes in each sub-bone’s controller is completely independent of the other sub-bones.
3. **A3**: Each sub-bone’s controller may use any animation technique to compute its transform including key-frame interpolation, procedure or motion capture.
4. **L3**: Each controller acts independently and can therefore be altered or even enabled and disabled independently of all the other processes.

The new framework also has several additional advantages:

1. It uses the existing structure of the avatar’s animation hierarchy, requiring no extra coding of layered controllers, or of process management systems in the sentence generator.
2. The computational burden is no greater than it would be with other options. No matter the system, the four contributing processes must be evaluated and their transformation matrices multiplied.
3. Each of the four sub-bones can manage its own linguistic process and will not interfere with the computation of neighboring processes.
4. As many bones as needed can be added per articulation in the avatar. The extra memory and data structure overhead for a bone is minimal.

One concern that may arise here is the question of how such an organization would affect the skin of the model. With increased numbers of bones in an avatar that affect the skin, there is increasing complexity in building the skin weighting factors, which determine how each bone affects the avatar’s skin.

However, if we look at the desired effect on the skin, we see that each of the processes that affect the sub-bones of waist must combine into a single transformation that will affect the skin of the model. All that happens is that one of the sub-bones can rotate the waist a little farther. The total effect on the skin will be the same. The proposed framework here assumes that these all contribute to one combined effect on both the torso and the model’s skin. As long as the most distal sub-bone in the waist controls the skin, the avatar will deform as expected. Thus no additional complexity in computing the deformation is introduced.

6 Conclusion and Future Work

The new framework presented in this paper allows for a flexible specification of many processes in sign simultaneously that all can influence any part of the anatomy. This is accomplished by replicating bones on a per-process basis so that each process can control its own bone independent of the others, thus allowing not only independent timing, but possibly completely different animation procedures or data controlling each. As long as

the interactions between the bones is minimal, the animation hierarchy will properly combine the process effects by multiplying their resulting transformation matrices.

Moving forward, there are a limited number of cases where interactions among tracks would be desirable. Consider situations where the transformation in one process will be influenced by the transformation in another. For example, the IK system on an arm may need to take into consideration some processes on the arm and torso and ignore others. Another example is the eyebrow motion described in [16]. In her study she found that in some cases the presence of one process can alter the range of motion in co-occurring processes. Such interactions will be addressed in a follow-up study.

References

1. Signed Language vs. Sign Language, Canadian Cultural Society of the Deaf. <http://www.deafculturecentre.ca/Public/Default.aspx?I=569&n>. Accessed 17 Jan 2017
2. Valli, C., Lucas, C.: *Linguistics of American Sign Language: An Introduction*. Gallaudet University Press, Washington D.C. (2000)
3. Elliott, R., Glauert, J.R., Kennaway, R., Marshall, I., Safar, E.: Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. *Univ. Access Inf. Soc.* **6**(4), 375–391 (2008)
4. Kipp, M., Heloir, A., Nguyen, Q.: Sign language avatars: animation and comprehensibility. In: *International Workshop on Intelligent Virtual Agents* (2011)
5. Crasborn, O., Sloetjes, H.: Enhanced ELAN functionality for sign language corpora. In: *Proceedings of the 3rd Workshop on the Representation and Processing of Sign Languages: Construction and Exploitation of Sign Language Corpora* (2008)
6. Hanke, T., Storz, J.: iLex – a database tool for integrating sign language corpus linguistics and sign language lexicography. In: *Workshop on the Representation and Processing of Sign Language, at the Sixth International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, Morocco (2008)
7. Kipp, M.: Multimedia annotation, querying and analysis in ANVIL. In: *Multimedia Information Extraction*, vol. 19 (2010)
8. Hanke, T.: Lexical sign language resources: synergies between empirical work and automatic language generation. In: *Language Resources and Evaluation Conference (LREC)*, Libson (2004)
9. Thorpe, D.: *Certified ASL Interpreter*. Personal Communication, Chicago (2016)
10. Lidell, S., Johnson, R.: American sign language: the phonological base. *Sign Lang. Stud.* **64**(1), 195–277 (1989)
11. Wilbur, R.: Phonological and prosodic layering of nonmanuals in American sign language. In: Emmorey, K., Lane, H.L., Bellugi, U., Klima, E. (eds.) *The Signs of Language Revisited: Festschrift for Ursula Bellugi and Edward Klima*, pp. 213–241 (2000)
12. Morgan, G.: Event packaging in British Sign Language discourse. In: Wiinston, E. (ed.) *Storytelling & Conversation: Discourse in Deaf Communities*, pp. 27–58. Gallaudet University Press, Washington, DC (1999)
13. Schick, B.: Classifier predicates in American Sign Language. *Int. J. Sign Linguist.* **1**(1), 15–40 (1990)
14. Filhol, M.: «Grammaire réursive non linéaire pour les langues des signes.» chez Conférence sur le Traitement Automatique des Langues Naturelles, Marseille (2014)

15. Wolfe, R., McDonald, J., Moncrief, R., Baowidan, S., Stumbo, M.: Inferring biomechanical kinematics from linguistic data: a case study for role shift. In: Symposium on Sign Language Translation and Avatar Technology (SLTAT), Paris, France (2015)
16. Weast, T.P.: Questions in American Sign Language: A quantitative analysis of raised and lowered eyebrows, ProQuest (2008)
17. Wolfe, R., Cook, P., McDonald, J., Schnepf, J.: Linguistics as structure in computer animation: Toward a more effective synthesis of brow motion in American Sign Language. *Nonmanuals Sign Lang. Spec. Issue Sign Lang. Linguist.* **14**(1), 179–199 (2011)
18. Lasseter, J.: Director, Toy Story. [Film]. Walt Disney Pictures/Pixar Studios, USA (1995)
19. Lee, J., Lee, K.H.: Precomputing avatar behavior from human motion data. *Graph. Models* **68**(2), 158–174 (2006)
20. Wolfe, R., McDonald, J., Schnepf, J.: An Avatar to Depict Sign Language: Building from Reusable Hand Animation, Berlin, Germany, 10–11 January 2011
21. Gibet, S., Courty, N., Duarte, K., Le Naour, T.: The SignCom system for data-driven animation of interactive virtual signers: methodology and evaluation. *ACM Trans. Interact. Intell. Syst. (TiiS)* **1**(1), 1–26 (2011)
22. Jennings, V., Elliott, R., Kennaway, R.: Requirements for a signing avatar. In: Workshop on Copora and Sign Language Technologies (CSLT), LREC, Malta (2010)
23. McDonald, J., Wolfe, R., Wilbur, R., Moncrief, R., Malaia, E., Fujimoto, S., Baowidan, S., Stec, J.: A new tool to facilitate prosodic analysis of motion capture data and a data-driven technique for the improvement of avatar motion. In: Lanugage Resources Evaluation Conference (LREC), Portoroz (2016)
24. Lefebvre-Albaret, F., Gibet, S., Turki, A., Hamon, L., Brun, R.: Overview of the Sign3D Project High-fidelity 3D recording, indexing and editing of French Sign Language content. In: Third International Symposium on Sign Language Translation and Avatar Technology (SLTAT), Chicago (2013)
25. McDonald, J., Wolfe, R., Schnepf, J., Hochgesang, J., Jamrozik, D., Stumbo, M., Berke, L., Bialek, M., Thomas, F.: An automated technique for real-time production of lifelike animations of American Sign Language, pp. 1–16 (2015)
26. Parent, R.: *Computer Animation: Algorithms and Techniques*. Newnes, Boston (2012)
27. Autodesk, Animation Layers (Layer Controller) Documentation. 15 November 2016. <https://knowledge.autodesk.com/support/3ds-max/learn-explore/caas/CloudHelp/cloudhelp/2017/ENU/3DSMax/files/GUID-ACAD407C-B079-4DAB-9DD0-6C8DC22389EF-htm.html>. Accessed 19 Jan 2017