Grazer Linguistische Studien 92 (Frühling 2021) S.168-228
DOI:***

A survey of facial nonmanual signals portrayed by avatar

Rosalee Wolfe & John C. McDonald

School of Computing, DePaul University Chicago

Abstract. Sign language avatars, also called signing avatars, are animated characters that display signed utterances on a digital display without using digitized recordings of human signers. These three-dimensional (3D) animated figures have the potential to increase Deaf accessibility to many of the resources in the hearing world. Before this can become a reality, however, researchers will need to address the challenge of effectively and accurately portraying facial nonmanual signals. This requires a thorough grounding in sign linguistics and the alternatives for its representation, as well as computer animation techniques. In this article, a discussion of the findings in these three fields forms the basis for analyzing past advances and current challenges in the portrayal of facial nonmanual signals by signing avatar.

Keywords. sign language, nonmanuals, avatar technology, sign language avatars

1. Introduction

An *avatar* is the computer portrayal of a person or a figure within an interactive user experience, and mention of avatars appeared in online games as early as 1979 (Caftanatov, 2018). Since then advances in computer graphics have facilitated ever more realistic figures in such applications as embodied conversational agents (Cassell, Vilhjálmsson, & Bickmore, 2004), in virtual worlds (Bartle, 2004) and the latest video games. For the past 20 years, researchers have developed avatars to portray sign languages. The goal of these *signing avatars* is to display signed languages as 3D animation, in lieu of displaying video recordings of human signers. The appeal of signing avatars is in their flexibility and consistency. It is easier to change or add a sign when using an avatar than it is to change or add a sign to a previously recorded video. Further, when a project requires repeated production sessions over a period of several weeks or months, it is easier to maintain presentation consistency in the lighting, camera angle, clothing and hair length of an avatar than it is when recording human signers (Wolfe, et al., 2016).

Signing avatars are currently in use to support dictionaries in such diverse sign languages as Czech, Indian, Italian, Saudi, French and American Sign Languages (Ahmad, et al., 2012) (Diwakar & Basu, 2008) (Elhadj, Zemirli, & Al-faraj, 2012) (Krňoul, Kanis, Železnỳ, & Müller, 2007) (Segouat, et al., 2008) (Vesel & Robillard, 2013) and Deaf education in many countries (Adamo-Villani & Anasingaraju, 2017) (Chiriac, Stoicu-Tivadar, & Podoleanu, 2015) (De Martino, et al., 2017) (Ferreira & García, 2018) (Jemni & Elghoul, 2008) (Krňoul, Kanis, Železnỳ, & Müller, 2007) (Shohieb, 2019) (Verlinden, Zwitserlood, & Frowein, 2005). Other applications that use sign language avatars include providing practice tools for interpreter education (Jamrozik, Davidson, McDonald, & Wolfe, 2010) (Toro, McDonald, & Wolfe, 2014), improving accessibility to conventional public address systems found at transportation hubs (Lombardo, Battaglino, Damiano, & Nunnari, 2011) (Ebling & Glauert, 2013), and generating synthetic data to train sign recognition systems (Brumm, Johnson, Hanke, Grigat, & Wolfe, 2018).

Signing avatars are a necessary component of automatic translation systems. For situations where the interaction is highly predictable but where an interpreter will never be available, a signing avatar can work as a component of an automatic translation system to provide rudimentary communication. Prototypes have translated weather reports (Grieve-Smith, 1999) (Verlinden, Zwitserlood, & Frowein, 2005), facilitated interactions with a post office clerk (Cox, et al., 2003) and airport security personnel (Furst, Alkoby, Lancaster, McDonald, & Wolfe, 2002), and have created Deaf-accessible public address systems (Ebling & Glauert, 2013).

Although avatar technology has improved significantly in recent years, there are still open questions about how best to display the linguistic and pragmatic information that occurs on a signer's face. The focus of this paper is a discussion of the potential of linguistics and computer graphics working together to portray facial nonmanual signals effectively through signing avatar technology. The remainder of this paper is organized into five sections. The first is a brief review of linguistic discoveries of the purpose and properties of facial nonmanual signals, and the second is a visual recounting of the developments in computer generated imagery (CGI) that make possible the computer display of signing. The third section is a discussion of representation systems that can direct an avatar's face to produce nonmanual signals, and the fourth section contains a description and analysis of past innovations and current efforts in signing avatar research, with a focus on facial nonmanual signals. The last section is a conclusion discussing promising research paths for future work.

2. A brief discussion of sign linguistics

It is surely appropriate to begin with a discussion of sign language linguistics, because it informs the development of signing avatars. The inception of sign language linguistics began with a focus on the manual aspects of the language. William Stokoe's pioneering work (1960) used the terms *tab*, *dez*, and *sig* to describe the position, handshape and movement of the hands while producing a sign. However, even in this early work, he observed that the production of a yes-no question involved a slight opening of the signer's face, "that is, his eyebrows raise and his chin lowers." (Stokoe, 1960, p. 62)

A little over fifteen years later, Baker (1976) affirmed that a signer's face does much more than communicate affect. She observed how affect can alter the form of a syntactic signal but even when altered, the signal is still recognizable (Baker-Shenk, 1985). Her studies of signers using American Sign Language (ASL) demonstrated how different combinations of facial and head movements communicate syntactic information including yes-no, wh- and rhetorical questions as well as topic marking. She noted that these syntactic constituents tend to appear on the upper half of the face. (Baker-Shenk, 1983). Signers use eyeblinks to mark constituent boundaries within discourse (Baker, 1978). In subsequent years, researchers continued to study eyeblinks to find that they can mark syntactic and prosodic phrases as well as discourse and narrative units (Wilbur R. , 1994). In addition, eye gaze can communicate syntactic agreement through the marking of referents (Bahan, 1997) (Bellugi & Fischer, 1972).

Baker also observed that facial behaviors on the lower face tend to modify the meaning of individual signs or phrases and communicate adjectival or adverbial information. For example, the nonmanual of pursed lips (Fehler: Verweis nicht gefunden) indicates that a surface is especially smooth (Wilbur R. B., 2009), and the presence of the nonmanual 'th' changes the lexical item LATE to NOT-YET (Reilly & Anderson, 2002).

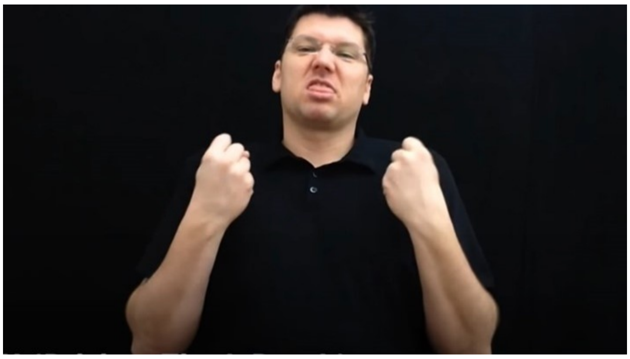
(Figure 1)

(Figure 2)

Since signed languages are not universal and vary by geographic location, it should come as no surprise that facial nonmanual signals also vary. One facial activity that that is acknowledged as part of many sign languages is *mouthing,* which is derived from words in the surrounding spoken language. Mouthing has been acknowledged as part of sign languages found in the UK, Sweden, the Netherlands (Crasborn, Van Der Kooij, Waters, Woll, & Mesch, 2008), Germany (Elliott & Jacobs, 2013), France (Sallandre, 2007), and the German-speaking regions of Switzerland (Braem, 2001). In contrast, mouthing in ASL is not considered part of the language and arises as a result of (English) language contact (Lucas & Valli, 1992), but more recent research has suggested that its usage is dependent on the grammatical category of the manual sign it accompanies (Nadolske & Rosenstock, 2007).

Fig. 1: The nonmanual 'pursed lips' in ASL (Shumaker, 2016)

Fig. 2 The nonmanual 'th' in ASL



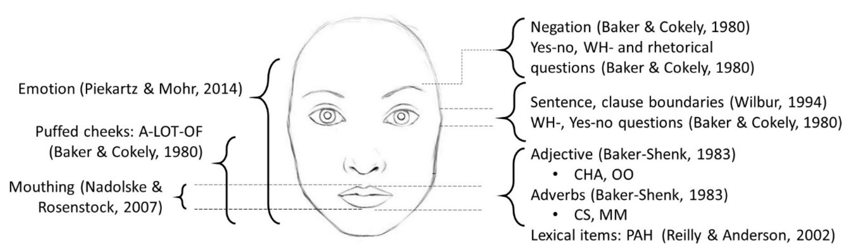
This short survey of discoveries into the processes involving facial expression in sign languages has demonstrated that facial expressions can convey information at all linguistic levels, phonemic to the prosodic, as well as communicating pragmatic information. This concludes our initial discussion of linguistic background. The next section will discuss the history of the technology necessary to portray a human face as computer animation. As a diagram that attempts to summarize the demands that linguistic and paralinguistic processes make of facial features. Please note that this diagram focuses on the facial articulators only, thus it omits the nonmanual articulators of the body (lean, shift), head (tilt, nod, shake, turn), teeth and tongue. (Baker & Cokely, 1980).

Figure 3

Conversely, every feature of the face communicates at least one and typically multiple functions in the structure of sign languages as well as communicating pragmatic information. For example, the brows can communicate both affect as well as syntactic information. Liddell (1980) noted that restrictive relative clauses involve not only the head and the upper part of the face (brow raise), but also a raised upper lip on the lower part of the face. Negation can involve a head shake, a lowering of the brow (upper face) and a raised upper lip (lower face). Several studies argue that adjectival and adverbial nonmanual modifiers occur on the lower part of the face. However, as can be seen in Figure 1 the nonmanual OO (smooth) has been modified by a brow lowering and an eye squint to intensify the meaning to be "very smooth". The intensifier, functioning as an adverb, occurs on the upper half of the face.

Even though a linguistic event is categorized as occurring on a particular facial feature, other features can participate in its production, but to a lesser degree. ~~Fehler: Verweis nicht gefunden~~ Figure 4 demonstrates how the eye aperture and brow height can change when a signer produces the ASL adverbial modifier CS, and ~~Fehler: Verweis nicht gefunden~~ Figure 5 shows eye aperture and brow variation in the production of puffed cheeks. Although the most salient facial features involved in producing the adverbial modifier CS are the lips, other features, located on other parts of the face, also participate. We will discuss this topic again in Section 4.4, Prosody and hybrid systems.

Fig. 3 Linguistic, pragmatic influences on facial features in ASL.

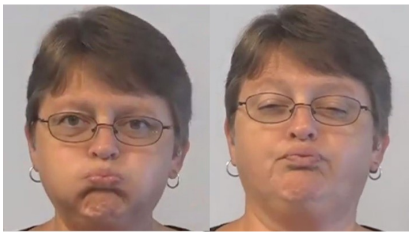


Move figure to top of this page.

This concludes our initial discussion of linguistic background. The next section will discuss the history of the technology necessary to portray a human face as computer animation.

Fig. 4 Neutral face and nonmanual marker CS. Note upper face participation (Foster, 2019)

Fig. 5 Puffed cheeks. Left: modifier A-LOT-Of, Right: *'A loooong time ago'* (Foster, 2019)



3. A visual history of computer animation

The technology of portraying facial expression via computer draws on many disciplines. Greek literature reveals that since ancient times, facial expressions have been known to add useful ancillary information to the words being spoken (Fergusson, 1940). In 1872, biologist Darwin proposed that facial expressions portraying six basic emotions were universal (Figure 6), but more recent studies suggest that facial expressions might be more culturally specific (Matsumoto & Hwang, 2011). In recent times, in-depth studies of facial behavior by psychologists have yielded a system of describing all visually discernible facial movement (Ekman, 1997). In fact, in her dissertation, Baker (1983) used Ekman's system in analyzing the nonmanual components of questions.

Insert "("

A contributing challenge to the difficulty of portraying a convincingly human face that moves in a lifelike manner is the complexity of facial muscles. Facial musculature varies greatly among individuals and some facial muscles are not present in all humans. However, all humans have the muscles required to produce the six universal emotions described by Darwin. (Waller, Cray Jr, &

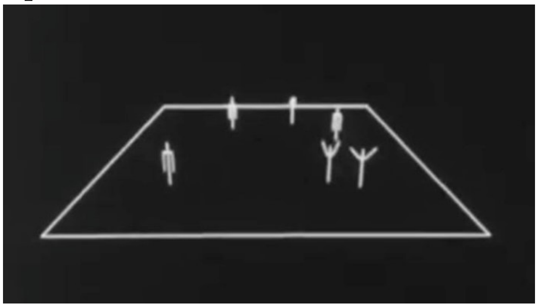
Burrows, 2008). As this section will show, current computer animation technology can portray these six emotional states quite effectively. Interestingly, research in sign language acquisition (McIntire & Reilly, 1988) indicates that although children consistently use facial expression to communicate emotion by the age of twelve months; they begin to acquire the grammatical facial behaviors later, at 24 months. Perhaps the greater length of time required to master the grammatical facial behaviors could indicate that these require the use of additional muscles.

Computer animation of humans began with very primitive figures with no recognizable faces. Similar to the development of sign language linguistics, the first computer animations of humans focused on the body. Creating depictions of the face only came later. This section presents a discussion of the development of computer animation or CGI (computer-generated imagery) over the past half century, with an emphasis on the realistic portrayal of humans as contrasted with cartoon figures. Although this section contains images representing the technological advancements, the reader is strongly encouraged to experience the corresponding animation as movement in time by following the links given in Table 1.

The first recorded computer-generated animation depicting human figures was Michael Noll's *A Computer Generated Ballet* (Noll A. , 1965), which consisted of white line segments drawn on a black background to create stick figures which moved about a rudimentary stage (Figure 7). Each arm of a dancer was represented as a single line segment, which alternately moved upwards and downwards. When dancers' arms were outstretched, it was possible to see that they could spin (rotate) as well as move (translate).

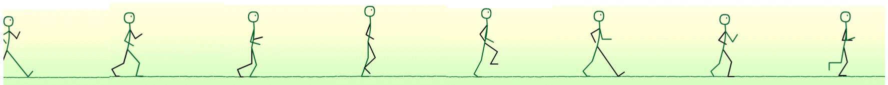
Fig. 6 Six basic emotions displayed facially. According to Darwin (Pierkatz & Mohr, 2014)

Fig. 7 A Computer Generated Ballet (Noll, 1965)



Although the visual depiction of the dancers would never be confused with a recording of actual ballet dancers in performance, this work represented a great step forward. To create the illusion of movement in a human's eye requires the quick display of a series of images at a rate of at least 20 frames (images) per second (Lext & Akenine-Möller, 2001). Figure 8 is an example set of frames depicting two steps in a walk. Drawing each of the frames by hand is a slow, painstaking process, even for simple stick figures, and a computer can speed what is otherwise a tedious and time-consuming task.

Fig. 8 Individual images of an animation. (Sanders, 2019)



Seven years after the introduction of *A Computer Generated Ballet*, Ed Catmull and Fred Parke developed a more realistic approach by using polygons, instead of line segments, to represent a hand. Figure 9 is a wireframe rendering of a hand, where the polygon outlines are clearly visible. Catmull created this model by making a plaster cast of his own hand, drawing polygons on the plaster, and measuring the position of each vertex in the polygon. Figure 10 shows the same hand, but in a smooth-shaded rendering. Because of the underlying polygon representation, the hand can be displayed as opaque in appearance, which is more realistic than a wireframe rendering. Parke then took on the challenge of creating a human face from polygons. Figure 11 and Figure 12 demonstrates the basic technique. This time, instead of making a plastic replica, Parke drew a grid pattern on his wife's face, and took two photographs – one from the front and one from the side. He used the photographs to measure the polygons.

Fig. 9 Wireframe rendering, from A Computer Animated Hand (Catmull & Parke, 1972)

Fig. 10 Smooth shading rendering, from A Computer Animated Hand (Catmull & Parke, 1972)

Fourteen years later, the 30-second commercial *Brilliance* (Figure 13) promoting canned foods featured a "sexy robot" character (Abel, 1984) who, despite her shiny metallic exterior, moves in a beautifully human-like manner. She shifts sinuously in her chair and reaches gracefully to light a candle. To recreate these movements, animators painted dots onto a human model who carried out the motions while the animators recorded her on film. This technique of recording human movements to apply to a computer figure is called *motion capture*. Although her body moved fluidly, her face remained motionless except for an occasional blink.

While *Brillance* featured an emotionless, idealized robot, *Tony de Peltrie* (Figure 14) features a human lounge pianist well past his prime who is wistfully remembering the successes of his youth (Bergeron, Lachapelle, Langlois, & Robidoux, 1985). This short film shows the first computer-animated human character to communicate emotion through facial expressions. While the character Tony imagines his past, his entire face reflects his thoughts and moods. One of the reasons that the character seemed more expressive was that the animators had better software that could create subtle, gentle movements which were an improvement over the awkward, mechanical movements in previous animations.

Fig. 13 Brilliance (Abel, 1984)

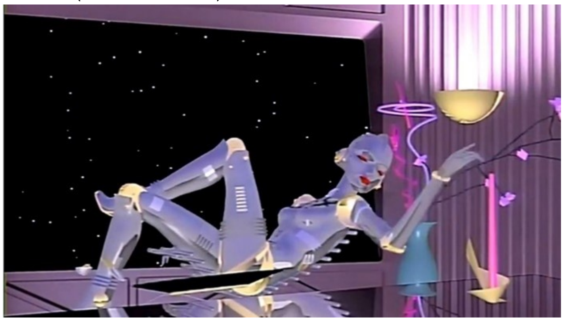


Fig. 14: Tony de Peltrie (Bergeron, Lachapelle, Langlois, & Robioux, 1985)

With subtlety came complexity. By now the polygon meshes had become sufficiently complex that specialized software became necessary to manipulate them. Researchers created a method called *rigging* to manipulate groups of polygons, instead of having to change each one. This is a two-step process. In the first step, animators create a virtual skeleton that is located inside the polygon mesh, just as a person's skeleton lies inside the body. In the second step, an animator attaches the polygons to various bones in the skeleton. Usually, a polygon is glued to the closest bone. Figure 15 shows the rigging of a virtual arm made from a polygon mesh. The skeleton consists of an upper arm bone and a forearm bone. These are shown in green. The polygons comprising the hand and the forearm are attached (skinned) to the forearm bone. They appear in fuchsia. The white polygons are attached to the upper arm.

The rigged arm is now ready to accept motion commands from an animator. Figure 16 shows the workflow of manipulating the arm in an animation. The left image in the figure shows the arm after an animator has selected the forearm bone, which is shown in green. The animator then rotates the bone, and all of the polygons attached to it rotate along with the bone. The result is displayed in the right image of Figure 16. The animator made one adjustment to the forearm bone instead of needing to adjust each individual polygon comprising the forearm. This technique saves an enormous amount of labor and makes possible the CGI imagery in today's animated movies.

Fig. 15 A rigged arm: a polygon mesh skinned to a set of bones.

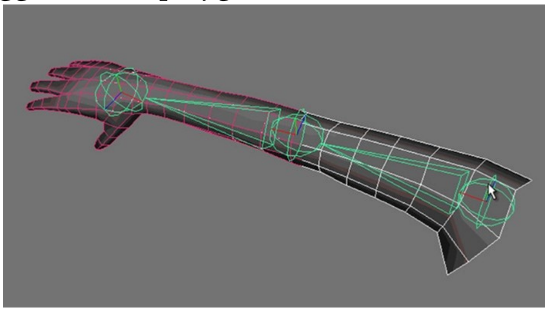
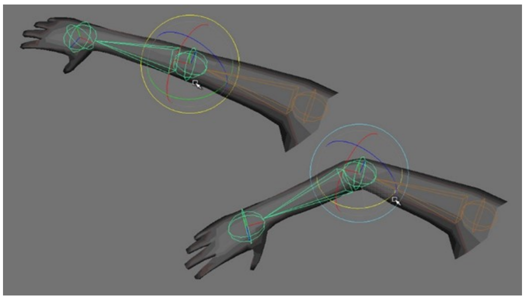


Fig. 16 Rotating the forearm



Luxo, Jr. (Figure 17) further expanded on the idea that motion conveys emotion through the use of the traditional Disney twelve principles of animation (Lasseter, 1986). Even though there are no human characters in the film, the two lamps effectively convey sadness, surprise, curiosity, parental resignation and joy. Lasseter stated that the *Tony de Peltrie* character inspired *Luxo Jr.*

After the success of *Luxo Jr.*, Lasseter and his team at Pixar undertook the ambitious project of animating a baby named "Billy" in the movie short, *Tin Toy* (Figure 18). It was enormously difficult to create (model) Billy's face and to animate it. Quoting Flip Phillips, a team member, "It just became an incredible burden" (Oleva, 2018). Early drafts of Billy's face looked more like that of a middle-aged man rather than a baby. The final version of Billy looked more like an actual baby and had 40 separate facial muscles controlling the polygon facial mesh, but his skin had the look of plastic.

Fig. 17 Luxo, Jr. (Lasseter, Luxo, Jr., 1986)

Fig. 18 Tin Toy (Lasseter, Tin Toy, 1988)

When it was initially released, researchers and animators alike hailed the film as ground breaking, but since then, viewers have often expressed an aversion to Billy, which is an example of the phenomenon known as the *uncanny valley* (Mori, 1970). Mori posited that a person's response to a robot would shift from empathy to revulsion if its face approached, but failed to attain, a lifelike appearance (Figure 19). He described this descent into eeriness as the uncanny valley. The effect is intensified if the robot appears to be moving (Figure 20). Thus, if the robot appeared eerie in a still pose, it will appear even eerier when it starts to move.

Fig. 19 The uncanny valley. After (Mori, 1970)

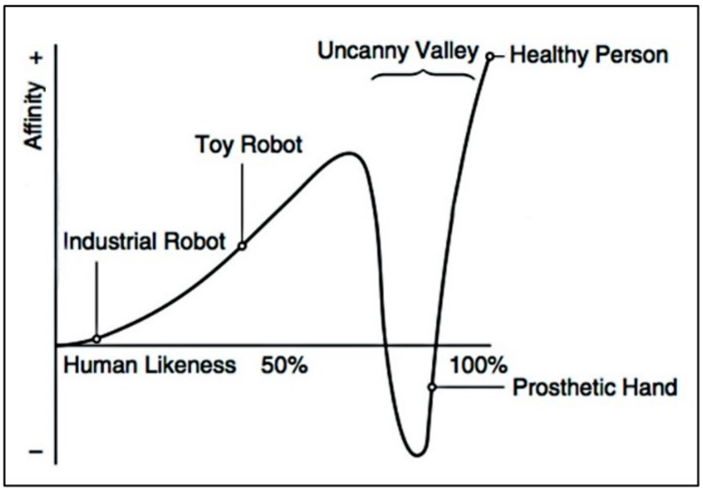
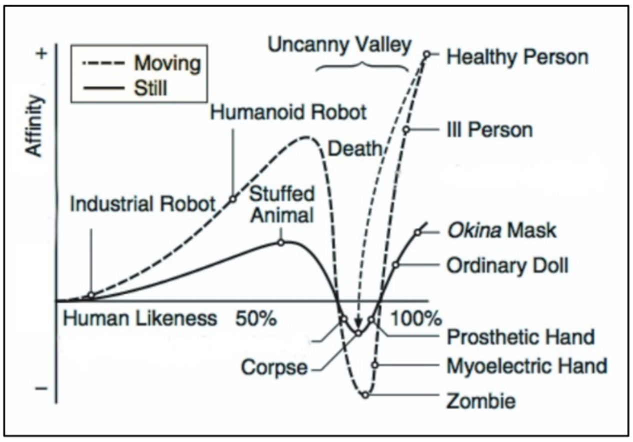


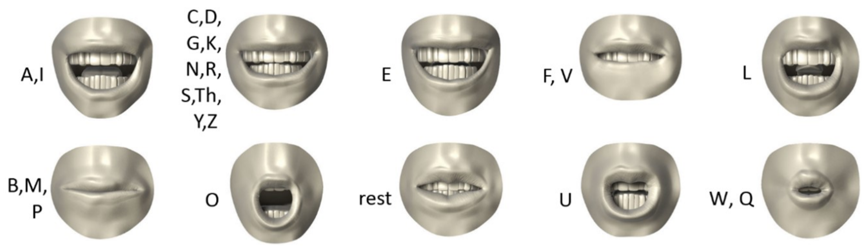
Fig. 20 Intensified by movement. After (Mori, 1970)



The movie The Abyss (Figure 21) actually took advantage of the uncanny valley by creating an aquatic alien being, “the Pseudopod,” which was expressly not human, and since it was made from water, did not have a human appearance except when it reconfigured itself to mirror the shape of a human face (Cameron, 1989). This movie also introduced the technique of morph target animation. In this approach, animators make copies of the neutral face, and then modify each copy to assume a different pose such as those depicting emotion, or mouth shapes for speech synchronization. Figure 22 shows a neutral face on the left, and two poses – one depicting a dropped jaw and another depicting puffed cheeks (Marmor, 2011). Animations are created by blending between different poses. These poses are also called blend shapes.

Automatic lip sync is possible using blend shapes. The term lip sync comes from the fact that animated characters who are speaking require lip motion that moves in synchrony with the soundtrack. Because some phonemes appear identical on the face even though they have different sounds, lip sync requires fewer mouth shapes than there are phonemes in a language (Figure 13).

Fig. 23 Viseme and corresponding phonemes commonly used in autoati lip sync of spoken english. After preston blair. (Martin, 2018)



Manually animating lip sync is tedious and time consuming which has motivated researchers to develop techniques to automate lip sync. Using a machine(computer)-readable format of the International Phonetic Alphabet (IPA) called SAMPA (Speech Assessment Methods Phonetic Alphabet), researchers have developed software to animate mouth movement in speaking characters (Xu, Feng, Marsella, & Shapiro, 2013).

The animators of the movie *Death Becomes Her* (Figure 24) successfully met the challenge of persuasively depicting human skin, including the twisted neck and stretched skin of Meryl Streep's character after she took a fall down a flight of stairs (Zemeckis, 1992).

In this film, the animators again used the uncanny valley to their advantage because they were creating creatures that were already dead.

The techniques of morph target animation and improved skin effects allowed directors to imagine new visual effects that were not previously possible and to also avoid costly reshoots. In the movie *Jurassic Park* (Figure 25), a stunt double portrayed the character Lex (Ariana Richards) in the scene where she falls through the ceiling but hangs on to climb upwards while escaping a velociraptor. In the shot, the stunt double looked up. The animators were able to replace the stunt double's face with Lex's face (Klassen, 2017).

Fig. 24 Death Becomes Her (Zemeckis, 1992)

Fig. 25 Jurrasic Park (Spielberg, 1993)

In the films previously mentioned, there had never been an extreme closeup on a computer-animated face, because of the dangers of the uncanny valley. However, the "Super Burly Brawl" scene in *The*

Matrix Revolutions (Figure 26) depicts exactly that, when the character Smith sustains a terrific face punch from Neo (Silver, 2003).

In subsequent years, software developers incorporated knowledge of physiological changes that humans undergo as they age coupled with advanced 3D scanning technology to create polygon meshes with ever finer detail. The result is that an actor can appear years younger in a flashback scene as Michael Douglas' character does in *Ant Man* (Figure 27) (Holmes, 2015) or that a new actor can assume a role when the original actor is no longer available as in *Rogue One: A Star Wars Story* (Emanuel, 2016). In *Rogue One,* the character of Grand Moff Tarkin is played by Guy Henry, with a synthetic copy of Cushing's face superimposed over Henry's (Figure 28).

Fig. 26 The Matrix Revolutions (Silver, 2003)

Fig. 27: Ant-Man (Holmes, 2015)

All of the facial animation techniques previously discussed use the same basic underlying technology, which is a three-dimensional polygon mesh. However, a recent technique uses only previously created video footage, which is two-dimensional. This is the *deepfake* approach (Wang, Gao, Tao, Yang, & Li, 2018). Suppose the desired effect is to substitute person A into a video of person B. In this approach, an extraction process identifies the faces in each frame (image) of video B and aligns them with the picture of person A. Then a training step allows a neural network to convert face A into face B. The final step "stitches" an altered version of picture A into each frame of video B. Figure 29 is a frame taken from a deepfake of John Oliver dancing instead of Jimmy Fallon. This example is particularly convincing because these two personalities have similar facial features. There are limitations to this approach. It requires an extensive amount of video data for the training step, and can be adversely affected by obstructions, such as the hands passing in front of the face.

Fig. 28: The making of rogue One: A star Wars Story (Emanuel, 2016)

Fig. 29: Deepfake of John Oliver dancing like Jimmy Fallon (WatchMojo.com, 2019)

Table 1 Additional resources for the animations discussed in the article

Year	Title	Available at	"Making of"
1965	A Computer Generated Ballet	https://youtu.be/uLU2hIV7n_I	https://youtu.be/-dXQIUEwEGQ
1972	Computer Animated Hand	https://youtu.be/fAhyBfLFyNA	https://tinyurl.com/wsp3eow
1984	Brilliance	https://youtu.be/7OGlLKzqHHk	https://youtu.be/Qnbpr0EB3nQ
1985	Tony de Peltrie	https://youtu.be/6GHJwBL1ySE	
1986	Luxo, Jr.	https://vk.com/video484769739_456239090	https://youtu.be/MJQRVKtwr70
1988	Tin Toy	https://youtu.be/-ejfNSPWMQE	https://youtu.be/bS2TXeXcx1I
1989	The Abyss	https://vk.com/video484769739_456239086 (excerpt)	https://youtu.be/gAFIUuFRkBA
1992	Death Becomes Her	https://youtu.be/p2hZia7wU-Q (excerpt)	https://youtu.be/xRNVHaaUAp8
1993	Jurassic Park	https://youtu.be/lbDFOAe_OXY (excerpt)	
2003	The Matrix Revolutions	https://youtu.be/Sh906GSszWk (excerpt)	https://www.youtube.com/watch?v=cnuZNapYNQM
2015	Ant Man	https://youtu.be/dr2rQnOQtfA (excerpt)	https://youtu.be/fT2pmyuoRyE
2016	Rogue One: A Star Wars Story	https://youtu.be/watch/1uKtUlPmhqw (excerpt)	https://youtu.be/HcXTPFgaa6E
2018	Excerpt from The Tonight Show	https://www.youtube.com/watch?v=q0wjV_Q-Lu8	

Broken link. Please use https://www.youtube.com/watch?v=D4NPQ8mfKU0

Broken link. Pls. use https://www.youtube.com/watch?v=wO7Vyhy0iXs

Broken link. Pls. use https://www.youtube.com/watch?v=SugS3936s5g

Delete entry

Delete entry

Delete entry

3.1 User interfaces for developing CGI

To produce movement, an animation artist has to understand the structure of the face being manipulated. The earliest interfaces were far more primitive than the ones in use today. They were simply scripts that specified poses and timing in a text file (Catmull, 1972). Needless to say, these were awkward and time-consuming to use. One had to type the information into a text file and submit it to the animation software to render, and it would be several hours before the finished image would be available.

As software and hardware improved, it became possible to make changes and see the effect immediately on a screen. At this point, it was possible to build user interfaces consisting of a set of controls that allows users to instantly see the results. According to Orvalho (2012) there are three forms of user interface: *window-based, 2D viewport* and *3D viewport.*

A window-based user interface (Figure 30) provides direct input of values, either by typing them or by adjusting a set of sliders. In a separate window, users can see the avatar's face and immediately view the effect of the changes in the values. In a 2D viewport interface, instead of typing numbers, the user manipulates a schematic representing the face, as seen on the right side of Figure 31. The user does not have to work directly with abstract numeric quantities but can change the 2D (flat) schematic and see the result on the face. There can be a bit of a disparity between the appearance of the flat schematic and the fully 3D face. However, in a 3D viewport interface, the controls are located directly on the avatar's face, which gives an artist the most realistic impression of how a manipulation will change the facial appearance (Figure 32).

Fig. 30 A window-based user interface (Schleifer, Scaduto-Mendola, Canetti, & Piretti, 2002)

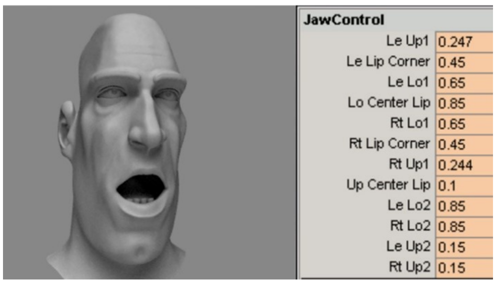
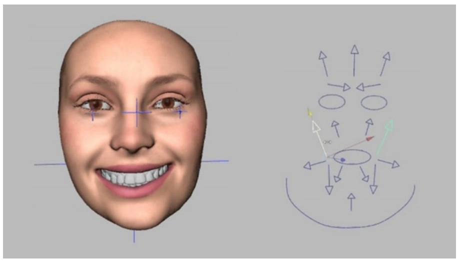


Fig. 31 A 2D viewport interface (Alexander, Rogers, Lambeth, Chiang, & Devebec, 2009)



3.2 A comparison of movie CGI to interactive graphics

The current state of the art of CGI yields photorealistic, convincingly life-like characters that now routinely appear in movies and television. However, creating animations of these characters is a time-consuming process. Each frame (image) of *Toy Story* required between two and thirteen hours to render (Wailes, 2014). At a rate of 24 frames per second, the render time for a second of animation required between 48 and 312 hours. Surprisingly, today's situation is not much better. Quoting Craig Good, a digital artist at Pixar and now retired (Good, 2016):

There's something I call The Law of Constancy of Pain: Back in 1983 it took between half an hour and around 8 hours to render a frame for one of our CG movies, such as André and Wally B. The average was probably a couple or three hours. Today, computers are literally millions of times more powerful. And guess how long it takes Pixar to render a frame.

Yup. Between half an hour and around 8 hours, with a typical average of a couple or three hours. Rendering time has stayed essentially flat for three decades. Why? Because the frames are now literally millions of times more complex and, apparently, humans are [still] willing to wait a few hours for a frame to render.

Such long rendering times are incompatible with interactive graphics, such as the ones found in video games. A video game must respond instantly to user commands. For this reason, the animation in video games and other interactive graphics applications must be greatly simplified in order to be as responsive to a user as possible. Due to the immense popularity of *The Matrix* series of films, the franchise created *The Matrix Online,* a massively multiplayer online role-playing game. The image in Figure 33 is a screen shot taken in a scene that evokes the “Super Burly Brawl” of *The Matrix Revolutions.* Comparing Figure 33 to Figure 26 reveals more primitive characters in a simpler environment, and the two characters are not touching each other. All of these reductions to the visual realism implies a reduction in rendering times which yields an increase in interactivity.

Fig. 32: A 3D viewport interface (Grubb, 2009)

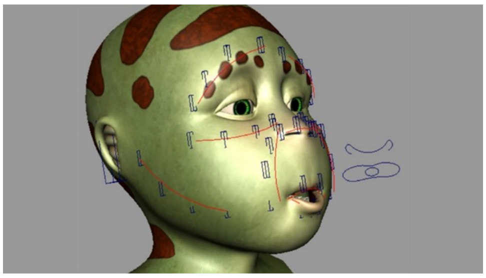


Fig. 33 Matrix Online (Royce, 2019)

Applying this insight to signing avatars, it is clear that an avatar in an interactive system will not have the same level of realism attainable when using CGI in a movie. It is possible to spend hours in creating an animation for film, but in an interactive system, the response must be immediate.

4. Representations supporting facial nonmanual signals

Once an avatar’s facial appearance has been defined and the user interface for the face has been determined, the next step is to create facial behaviors. This requires data that describes how the facial features should move. Once these motions are defined, the third and final step is displaying the avatar's face, on a computer screen.

There are three alternatives for acquiring the data to determine facial movement. The first encompasses traditional animation techniques; the second uses data from sign language notation and/or transcription systems, and the third is a hybrid approach.

4.1 Traditional animation

Traditional animation techniques include *keyframe animation* and motion capture. In keyframe animation, an artist sets up the starting and ending poses for any smooth transition. The leftmost and rightmost images in Figure 34 are examples of a key frames. An artist created these poses. The two images in between the two key frames are examples of *inbetween frames,* created by the computer (Burtnyk & Wein, 1976). Motion capture, as mentioned previously, can record the positions of an actor's features over time, as is seen in the left image of Figure 28.

The advantage of both of these traditional animation methods is that they can create high-quality, lifelike motion. However, there are disadvantages as well. They do not include any linguistic information. Data collected in this fashion requires additional information annotated by linguists in order to use the data for anything other than playback. Without linguistic intent, any newly formed sentences appear awkward and are difficult to read. Even though this data is not yet amenable to automatic linguistic analysis, there is a lively community of researchers examining this problem (Koller, Zargaran, Ney, & Bowden, 2018).

4.2 Graphics standards

A second disadvantage of traditional animation methods is the extremely limited quantity of data available in this form, which stems from the lack of cost-effective methods to acquire data quickly. A number of research institutions, governments, and industries realized that one way to address this problem is to share information. They would benefit if they could share the cost of acquiring and displaying animation by establishing open (published) standards for 3D graphics. The goal of *Web3D* is to develop and maintain royalty-free standards that can work together and run on all platforms and devices, including desktops, tablets and phones (Web 3D Consortium, 1999). This approach allows multiple software projects to utilize pre-existing animation services, thus freeing the majority of software developers from (re)writing graphics software that is necessary for their projects.

In 1995, VRML (Virtual Reality Markup Language) became the first web-based 3D format, created to support virtual worlds in interactive gaming. The current standard, X3D, subsumes VRML and provides other standards as well (Campbell, 2013). One of these is MPEG-4 H-Anim, the Humanoid Animation International Standard. This standard facilitates the display of 3D avatars in web browsers (such as Chrome or Safari) on any device.

The H-Anim standard provides for Body/Face Definition Parameters, which define the appearance of an avatar, and Body/Face Animation Parameters, which control the movement of the avatar. Figure 35 shows three 3D figures that are H-Anim compliant, and are displayable in any browser with the use of a plug-in. Figure 35 shows the H-Anim facial feature points, used to designate Face Definition Parameters. These parameters were used to define the two faces seen in Figure 37. Each of the emotions were created through a definition of FACS settings which were supported by H-Anim Facial Animation Parameters.

Fig. 34 Keyframes and in-betweens (Pluralsight Creative, 2013)

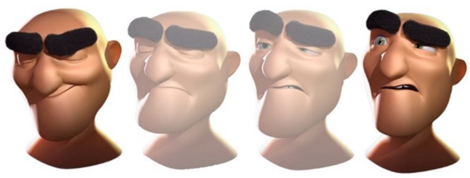


Fig. 35 Three figures defined by H-Anim Body Definition Parameters (Autodesk, 2019)

Fig. 36 H-Anim feature points, used in designating Face Definition Parameters

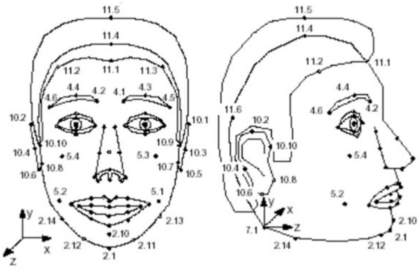
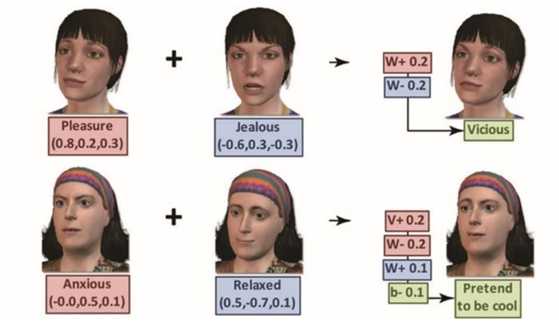


Figure 37 Using H-Anim feature points to support protraying FACS-based emotions



4.3 Sign language notation and annotation systems

An alternative source for data comes from previously established sign language notation systems. The earliest of these was Stokoe notation, the first phonemic notation for a sign language (Stokoe, 1960). Created specifically for ASL, it provided for the specification of location (tab), handshape (dez) and movement (sig), but not for nonmanual signals.

In contrast, HamNoSys (the Hamburg Notation System) is a phonetic notation that can specify handshape, palm orientation, location, movement and nonmanual signals (Hanke, 2004). It can accommodate a broad range of sign languages as was demonstrated by the Dicta-Sign project (Efthimiou, et al., 2010) which created corpora in British, German, French and Greek Sign Languages, all notated in HamNoSys. Because HamNoSys was not originally intended for computer processing, researchers created a new XML-compliant representation of HamNoSys, named SiGML (Signing Gesture Markup Language) which is easier to process by computer (Glauert & Elliott, 2011).

Although not strictly a notation system, SignWriting is a writing system designed for a non-specialist to describe a sign or signed sentence "for everyday purposes" (Kato, 2008). It is a pictorial writing system of characters that are abstract pictures of the hands, face, and body. Figure 39 demonstrates the ASL sign WHO in SignWriting. The brows symbols indicate that they are lowered, and the symbol for the mouth indicates that the mouth is puckered. The spatial arrangement of the symbols on the page is two-dimensional, rather than the linear form of Stokoe or HamNoSys notation. SignWriting is used in Deaf education in Germany and Brazil, and researchers have developed an XML-compliant version called SWML (Papadogiorgaki, Grammalidis, Makris, Sarris, & Strintzis, 2004).

Fig. 38 HAMBURG in DGS: sign sketch and HamNosys notation (Hanke, 2004)

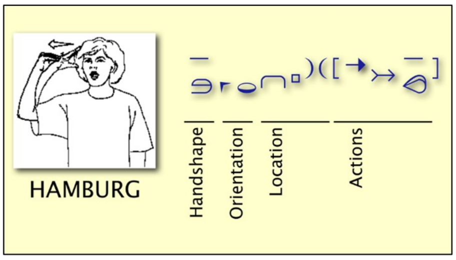
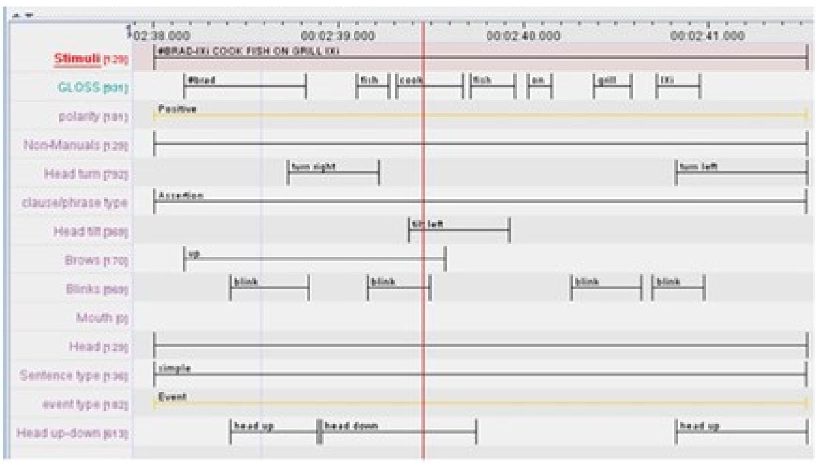


Fig. 39 ASL sign WHO in SignWriting (Sutton, 2013)

One of the advantages of notation systems is that they are amenable to automatic analysis through statistical tools which can support research projects. However, these notation systems can express only sparse information about the timing of the phonemes within a sign or within the prosody of a sentence. Further, linear notation systems do not model (specify) timing of co-occurrences effectively. From the earliest linguistic studies, researchers noted that there were multiple channels in signed languages. Baker-Shenk (1983) noted five channels – eyes, face, head, hands and arms – where behaviors could co-occur. In recent years, researchers have identified as many as fourteen channels (Wilbur R. B., 2009), most of which occur on the face. Indeed, quoting from Baker-Shenk, "Nonmanual signals are not composed of a single behavior, but a configuration of several behaviors" (1983, p. 103). Further, the timing of facial actions helps to distinguish nonmanual signals. Being able to produce the onset, intensity, and duration of the behavior in nonmanual signals is necessary for their clear communication. At the present time, this and their co-occurrence are difficult to express as notation.

Two useful annotation systems that can record co-occurrences are iLex (Hanke, 2002) and ELAN (Brugman, Russel, & Nijmegen, 2004). Both systems are time-based systems for annotating video. They allow for the definition of tiers that correspond to the channels being studied. Figure 40 is a partial screenshot from an ELAN file which contains multiple tiers representing various linguistic processes involving nonmanual signals. There are many instances of co-occurrence in this short example.

Fig. 40 An ELAN annotation with tiers (Benitez-Qiroz, Gököz, Wilbur, & Martinez, 2014)



4.4 Prosody and hybrid systems

Many of the previously mentioned representation methods focus on the lexical, morphological, and phonemic levels, but in recent years, more attention has been directed toward larger groupings of these elements into prosodic units. Prosody refers to the rhythmic phrasing, stress, and intonation that gives coherence to an utterance (Adamo-Villani & Wilbur, 2015). Knowledge of prosody is necessary for natural, easy-to-read generation of utterances signed by avatar. In the absence of prosody, the animated signing "will be as unacceptable and potentially as difficult to understand as robotic speech lacking cues to phrasing, stress, and intonation." (Adamo-Villani & Wilbur, 2015, p. 308)

Wilbur (2009) identified several prosodic elements that are predictable and can be automatically incorporated into automatic sign animation including pauses between phrasal constituents, noun or verb phrases, and sentence boundaries. Predictable nonmanual movements are head shaking throughout a negated phrase, brow movements marking questions, topics, and conditions, and eyeblinks to either mark the ends of higher prosodic phrases, or to provide emphasis. However, many facial nonmanual markers are not as predictable and must be recorded in a representation. For this, Adamo-Villani and Wilbur created the representation ASL-Pro (ASL with prosody).

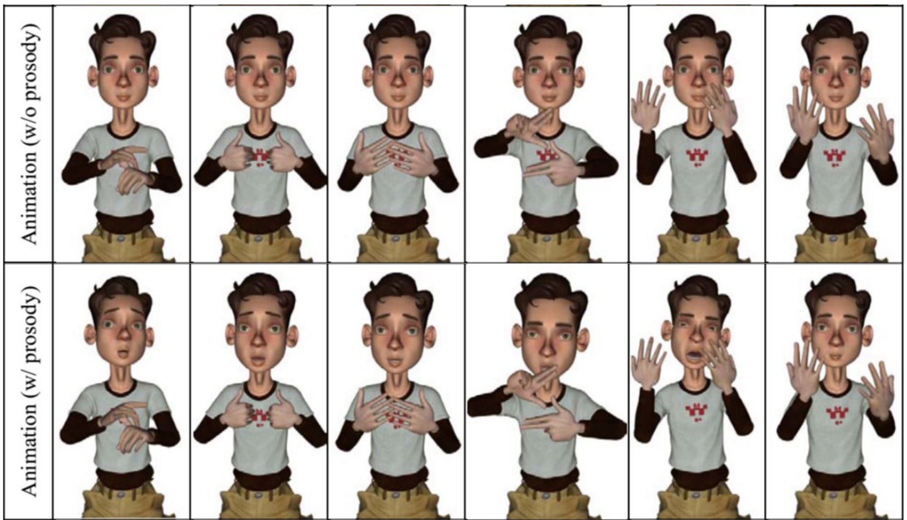
At first glance, the ASL-Pro looks nearly identical to ASL gloss notation, but there are some significant differences which makes it amenable to computer processing (Figure 41). There are no lines above sign phrases. A comma after the initial noun NATURE indicates a brow raise; the period at the end of the sentence will create a blink and phrase final lengthening in the animation. Additional prosodic events are indicated in square brackets.

Figure 42 shows an animated avatar signing without prosody, an animated avatar signing with prosody, and a timeline with layers of the representation. In the timeline, row A shows the gloss stream and Row C contains the English translation. Row B specifies the phrases, and rows D, E, and F specify the prosodic markers appearing on the brows, head, and mouth, respectively. Row G shows prosodic structure that can be predicted automatically, and Row H shows the resulting effects. The result signing on the avatar is clear and easy to read.

Fig. 41 A comparison of ASL-Pro notation to ASL gloss notation (Adamo-Villani & Wilbur, 2015)

English	Nature can be unbelievably powerful.
ASL gloss	br bl lb, hu, bf bl lf bl NATURE SELF-1 WOW AWESOME POWER . br = brow raise; bl = blink; lb = lean back; hu=head up; bf=brow furrow; lf = lean forward
ASL-Pro notation	NATURE, SELF-1 WOW[emph] AWESOME POWER[focus].

Fig. 42 ASL with, and without prosody. (Adamo-Villani & Wilbur, 2015)



A	NATURE.	HAVE	WHAT.	HURRICANE	FOREST-FIRE	return-to-neutral.
B	nature-topic	have	what-end phrase	hurricane	forest fires	end of clause
C	"What nature has is hurricanes and forest fires."					
D	up-topic	up-wh-cleft	→	neutral	down-emphasis	neutral
E	back-topic	forward	→	down	back-emphasis	neutral
F	'na'	'ha……………….ve'			'wow'-emphasis	neutral
G	topic	wh-cleft	→	list item$_1$ … last list item$_4$-emphasis		
H	comma phrasal lengthening slower speed, add pause			period sentence final lengthening slower speed, add longest pause		

Another way to represent prosody is through a hybrid approach. Hybrid representation systems incorporate prosody by combining the abstraction of linguistic labelling with the precision of traditional animation techniques. This approach can represent the linguistic content, as well as the biomechanical details of the avatar.

One such approach achieves very human-like motion by incorporating motion capture information which records the full range of fine biomechanical detail. While motion capture seems to be a natural way to transfer signing motion to an avatar, it has many of the same issues as video in terms of generating new utterances that were not previously recorded. One approach to utilizing motion capture information is to cut it into short clips that can then be chained together to create the desired utterance (Gibet, Courty, Duarte, & Naour, 2011). Ongoing research is addressing issues that hamper stitching clips such as the identification of differing starting and ending positions of linguistic events which can complicate the movement between clips. Additional challenges include the identification and removal of prosodic events which are present in a recorded clip but are not appropriate for the new utterance being generated.

The need for representing prosody is where hybrid approaches have become important since one of the key areas where prosody affects the production of utterances is in facial non-manual signals. In this case, the motion capture data is restricted to the body and a system of blend-shapes, scheduled by linguistic information, is used on the face (Gibet, Lefebvre-Albaret, Hamon, Brun, & Turki, 2016).

A second hybrid approach (Filhol & McDonald, 2018) relies on a mixture of traditional and procedural animation to build the basic elements for discourse produced by avatar. These include

1. Pre-animated sequences from artists for highly repeatable portions of sign that rarely change form such as lexical items,
2. Non-manual sequences such as movements of the head or torso that communicate the structure of sentences,
3. Non-manual facial expressions whose intensity can be controlled and mixed by a collection of animation curves,
4. Classifier and size and shape constructs which can be laid out procedurally.

In this second approach a hierarchical description called AZee provides the coordination and timing of these animation data to create co-occurrences (Filhol & Hadjadj, 2016). The hierarchical nature of the description gives the animation system the flexibility to decide how to compose the animation so as to best leverage the natural motion defined by artists. The goal of this system is to be able to build directly from linguistic descriptions while relying on artist intuition and mathematical procedures to add the needed human naturalness. When the animation system works with larger segments of the discourse, rather than sequences of individual phonemes, it is able to achieve more natural animation, due in large part to the inclusion of prosody, either explicitly in the linguistic description, or implicitly through the animator's eye.

The power of this particular hybrid system arises from the fact that both the avatar and the linguistic representation allow multiple processes to affect different parts of the body at the same time and allow asynchronous timing of processes. While linguistic systems often make simplifying assumptions that discretize the human body into zones for the communication of manual and non-manual signals, studies like (Weast, 2008) indicate that these signals will actually mix legibly on features such as the eyebrows for communicating yes-no or wh-questions and emotional states such as joy or anger.

As noted in Section 2, natural human motion that carries out facial nonmanual signals can affect multiple facial features due to the complex interconnectedness of the human musculature. For example, when signing the ASL facial nonmanual signal CS, more than the lips will move. This nonmanual will also affect the cheek muscles as well as eye aperture. Further, the timing of facial feature movements may not be synchronous as each may have different onsets and durations. The ability to layer and schedule such processes with asynchronous timing greatly facilitates the inclusion and coordination of prosodic elements which are rarely perfectly synchronized on lexical boundaries.

5. Progress in developing avatar systems

The first signing avatar efforts focused on the manual channels of sign language. The SignSynth project (Grieve-Smith, 1999) used a text representation of Stokoe notation and converted it into VRML which could be displayed in any web browser that was X3D capable. The SignSynth avatar is in the middle image of Figure 43.

Other early effort efforts did not work directly with VRML but made use of general-purpose avatars that had been created for other applications. Zhao's (Zhao, et al., 2000) efforts utilized *Jack*, a human simulation system originally designed for ergonomics research. The underlying sign representation was based on Laban dance notation. Another group utilized a pre-existing BAP (H-Anim Body Animation Parameters) player to display SignWriting (Papadogiorgaki, Grammalidis, Makris, Sarris, & Strintzis, 2004). Envisioned applications for these early avatars included interpreting broadcast news, particularly weather reports, dictionaries, and text-to-sign translators. In Fehler: Verweis nicht gefunden the Jack figure is on the left and the BAP player on the right.

Some of the earliest nonmanuals implemented were eye gaze, blink, head tilt, and torso rotation (Karpouzis, Caridakis, Fotinea, & Efthimiou, 2007). From a technical viewpoint these are easier to implement than are individual mouth movements. Because modeling the movement of the forehead is more constrained than the lower portion of the face, development of believable brow movement occurred before the development of lower face movement (Craft, et al., 2000). See Figure 44.

Fig. 43 Early avatars. From left to right: (Zhao, et al., 2000), (Grieve-Smith, 1999), (Papadogiorgaki, Grammalidis, Makris, Sarris, & Strintzis, 2004)

Fig. 44 Early efforts at facial nonmanual signals. Left: (Karpouzis, Cardakis, Fotinea, & Efthimiou, 2007) Right: (Craft, et al., 2000)

The first commercially available signing avatar (Figure 45) in the United States that provided basic facial expressions was Vcom3D (Sims & Silverglate, 2002). It implemented facial expressions as a set of blend shapes, sequentially displayed. The avatar was H-Anim compliant and thus capable of working in a web browser with an additional plug-in. Vcom3D also offered a sign language editor, which enabled research efforts to create a science dictionary (Vesel, 2005), educational modules for teaching mathematics (Adamo-Villani, Doublestein, & Martin, 2004) and STEM topics (Andrei, Osborne, & Smith, 2013).

Other research efforts explored the utility of *embodied agents* which were avatar systems built to speed up the time-consuming process of manual character animation (Heloir & Kipp, 2009). EMBR (Embodied Agents Behavior Realizer) a real-time character animation engine (Figure 45), served as a test bed for the perception tests of sign language utterances produced by avatar (Huenerfauth & Kacorri, 2015) (Kacorri & Huenerfauth, 2016).

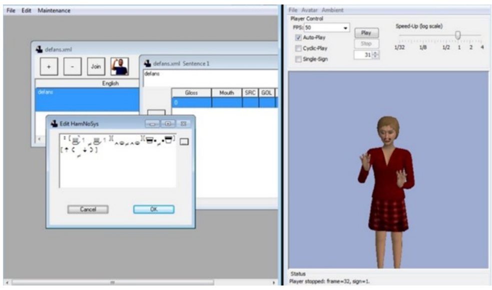
In the past two decades, the European Union (E.U.) supported a series of research projects that created and refined a set of avatars with the goal of improving access of information and services to the Deaf community in their preferred sign language. Beginning in 2000 with the ViSiCAST project (Elliott R. , Glauert, Kennaway, & Marshall, 2000) researchers built on previously special-purpose systems to create a web-compliant avatar capable of displaying manual (Kennaway R. , 2001) and nonmanual signals (Elliott, Glauert, & Kennaway, 2004) in any sign language, not just in one specific language. The initial version of the avatar used HamNoSys as its underlying representation, which is a language-neutral phonetic system rather than a language-specific phonemic system. In subsequent versions, researchers created SIGML, which is HamNoSys in an XML-compliant input form, and more amenable to computer processing. SiGML includes the facial expressions specified in HamNoSys 4.

The second project in the series, eSIGN, included an editor (Figure 46) that facilitated the creation of a sign lexicon in any signed language by specifying the corresponding HamNoSys (Hanke, Popescu, & Schmaling, 2003). The eSIGN editor also facilitates the creation of signed sequences by selecting signs from any lexicon that was previously created with the editor. Users can view the resulting animation on an avatar and choose to edit the sequence by modifying the morphology or phonetics as necessary. The eSIGN project created a variety of avatars (Figure 47) including Anna, an avatar whose proportions are carefully modeled to match those of a human as closely as possible (Figure 46).

The third project in the series, Dicta-Sign (Matthes, et al., 2012), used the tools developed in the eSIGN project to produce parallel sign corpora in British, French, German and Greek Sign Languages (Efthimiou, et al., 2010). During the Dicta-Sign Project, researchers made significant refinements to the facial nonmanual signals (Jennings, Elliott, Kennaway, & Glauert, 2010) including *masked blend shapes* to implement the nonmanual signals specified in SiGML as well as mouthing specified as SAMPA. Unlike a conventional blend shape, a masked blend shape only affects part of the face. Figure 48 demonstrates how a masked blend shape is only affecting the upper lip of the avatar.

Fig. 45 Examples of Vcom3D (Sims & Silverglate, 2002) and EMBR avatars (Heloir & Kipp, 2009)

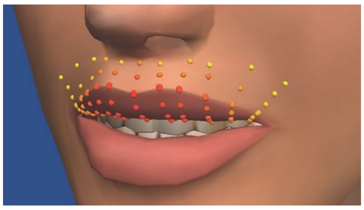
Fig. 46 eSIGN editor with Anna



In another project within Dicta-Sign, Elliott et al. (2010) and Glauert et al. (2011) spearheaded efforts to add timing information to the SIGML specification to create more finely grained avatar control with the goal of creating more legible mouth shapes, both for mouthing and for mouth gestures. The result was JASigning, signing avatar software that can display sign languages used in different countries. It accepts SIGML input which it converts to animation. Figure 49 shows examples of the refined mouth shapes.

Fig. 47 eSIGN project avatars: Siggi, Anna, Marc and Luna

Fig. 48 A masked blend shape for the upper lip



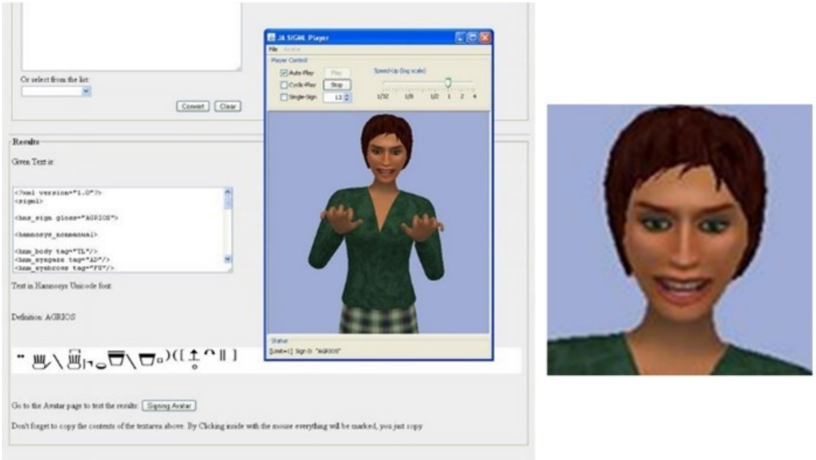
Incorrect font

The release of the tools created by the ViSiCAST, eSIGN and Dicta-Sign efforts enabled additional resource development. Krňoul et al. (2007) developed custom software that utilized the manual portion of the HamNoSys specification coupled with visemes to produce mouthings. Ebling (2016) used the JASigning avatar to display translations of train announcement previously given only as spoken statements over a loudspeaker in a train station.

The underlying technology developed in the E.U. projects facilitates the creation additional nonmanual signals which are then available for generating new signed utterances. Building on the tools developed by the E.U. projects, the research group at the Institute for Language and Speech Processing (ILSP) have created additional representation for lexical and syntactic features of Greek Sign Language, including nonmanual signals (Goulas, Fotinea, Efthimiou, & Pissaris, 2010) (Kouremenos, Fotinea, Efthimiou, & Ntalianis, 2010) and developed an online editor Sis-Builder designed to expand and maintain sign language resources (Efthimiou, et al., 2019). Sis-Builder can facilitate building new lemmas from HamNoSys, or building signed utterances using previous-defined lemmas for the avatar Anna to sign (Figure 50).

Fig. 49 Examples of mouth shapes from the Dicta-Sign Project

Fig. 50 Sis-Builder session. Anna is signing WIKD in Greek Sign Language



Masked blend shapes were also used in Mathsigner (Hayward, Adamo-Villani, & Lestina, 2010), to create interactive learning tools to improve the mathematics skills of Deaf children. Educators can create lessons that use an avatar for instruction. They can pick the facial expressions for the avatar from a palette of predefined options (Figure 51). The ATLAS project (Lombardo, Battaglino, Damiano, & Nunnari, 2011) also makes use of masked blend shapes in in their implementation of Donna (Figure 52), a virtual interpreter of Italian Sign Language (LIS).

Figure 51: Mathsigner editing window. User can choose masked blend shapes for individual features

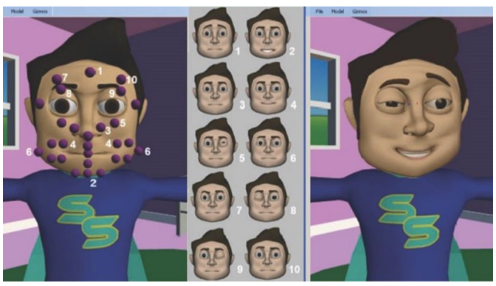
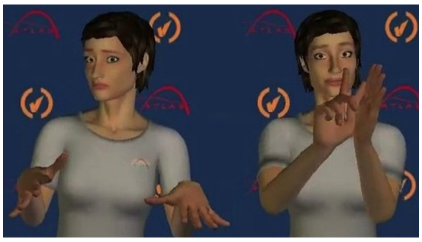


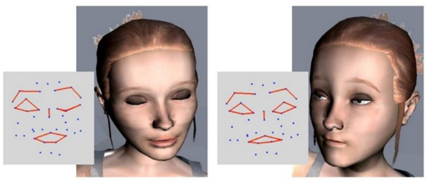
Fig. 52 Donna, a virtual interpreter for LIS



These systems using blend shapes all give a user the option of choosing a facial expression that is a static pose, and the animation engine automatically computes the transitions between the poses. This often leads to facial motion that looks stiff, awkward, or unnatural. In the SignCom project, Gibet et al. (2011) use motion capture of a native signer (Figure 53) to guide the selection of facial expressions. Figure 54 demonstrates how changes in the facial markers are create changes in the selection and influence of the blend shapes. The resulting movement is more lifelike, as it was based entirely on human movement.

Fig. 53 Motion capture setup, with markers on the face

Fig. 54 Sample facial positions, along with corresponding markers position



Although the majority of projects use HamNoSys as part of their sign representation system, two projects, one in Tunisia (Bouzid & Jemni, 2014) and the other in South Africa (Moemedi & Connan, 2010), use SignWriting instead. To implement facial poses specified via H-Anim Facial Animation Parameters (FAP), both efforts use a set of additional bones to change the shape of the facial polygon mesh. Sample mouth and eyebrow poses are shown inFigure 55 and Figure 56.

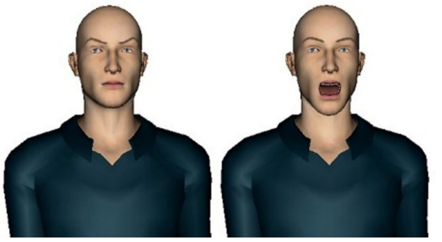
Insert space

SignWriting

eyebrow

Fig. 55 Mouth, exebrow poses supporting SignWritin (Bouzid & Jemni, 2914)

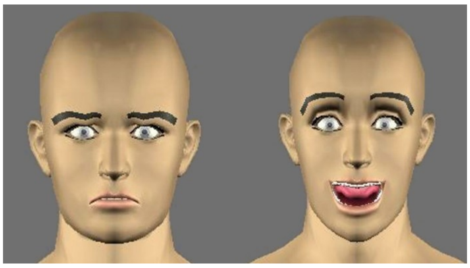
2014



eyebrow

SignWriting

Fig. 56 Mouth, exebrow poses supporting SignWritin (Moemedi & Connan, 2010)



As mentioned in Sections 2 and 4.4, the nature of co-occurring facial nonmanual signals will produce multiple linguistic events that simultaneously influence the face. Further, a signer's production of co-occurring nonmanual signals will often result in an individual facial feature participating in producing multiple linguistic events simultaneously. To quote Baker-Shenk (1983, p. 103), "consider the reports that many nonmanual signals are not composed of a single behavior, but of a configuration of several behaviors." For an extensive list of scenarios where this occurs in ASL, please see (Baker-Shenk, 1983).

To handle simultaneously-occurring events on the face, Gibet's approach (Gibet, Courty, Duarte, & Naour, 2011) collects corpora of annotated motion capture of facial markers, and selects the closest-matching blend shapes from a library created by artists/animators as demonstrated in Figure 54. The quality of the resulting animation depends on having a large amount of motion capture data in a corpus.

There is a challenge in making changes or corrections to the resulting animation because the motion is a combination of *all* of the nonmanual events. Suppose the avatar posed a yes/no question: "Did you see your mother recently?" In ASL, the yes/no-question marker would raise the avatar's brows and open the eye aperture. The CS adverbial would change the mouth shape and narrow the eye aperture. If a fluent ASL user would want to remove the adverb from the utterance, the only option would be to search the corpus again.

A second approach, introduced by Schnepp (2006) and generalized by McDonald et al. (2017) represents each linguistic and paralinguistic event as a separate stream (or tier) of co-occurring influences. These are combined on the avatar to portray all linguistic and paralinguistic activity. Figure 57 demonstrates the approach.

In the leftmost image, the avatar "Paula" is signing YOU as a statement in neutral demeanor. The next picture is a closeup of her face while signing the statement. In the middle picture, Paula is signing the interrogative YOU? Her head tilts forward, her brows rise and her eye apertures open wide. The next image is a closeup of Paula's face while posing the interrogative YOU?! in an angry manner. Her head is still tilted forward, but her eye apertures have narrowed a bit and furrows appear between her lowered brows. Compare this to the image on the far right which is a closeup of Paula's face while signing the statement YOU! in an angry manner. Her head is upright; her brows are lowered; furrows appear between her brows, and her eye apertures are tightly narrowed.

An advantage of this approach is that viewers can distinguish the co-occurring linguistic events in utterances signed by Paula (Schnepp, Wolfe, McDonald, & Toro, 2013), because it is easy to edit individual linguistic contributions to the facial features. Each stream also has its own timing information, independent from other streams which frees the nonmanual signals from starting at the same time as the onset of a lexical item.

Fig. 57 Left to right: Paula signing YOU in a statement with neutral demeanor. Closeups of a) statement, neutral demeanor b) interrogative c) interrogative, angry demeanor, d) statement, angry demeanor

In a recent trend, some research projects and commercial products have introduced avatars with a more cartoon-like appearance, including products by Hand Talk (Hand Talk, 2019), Huawei (Huawei, 2020),and Simax (Pauser, 2019) as well as Adamo-Villani's avatar Jason (2015). Hand Talk recently acquired ProDeaf, a research group focused on Portuguese-LIBRAS translation (Hand Talk, 2019) and introduced a new avatar Hugo, seen as the central character in Figure 58. Huawei's developers created their avatar Star as part of a literacy platform for Deaf children. Simax offers automatic translation of such content as websites and traffic news, primarily in Austrian Sign Language. Adamo-Villani created used the Jason avatar in her prosody studies mentioned previously. Figure 59 shows all three avatars: Simax, Star, and Jason.

Avatars created in the cartoon style have one strong advantage: they avoid the trough of the uncanny valley, and the viewer does not need to cope with the distraction of eeriness that would otherwise arise. The exaggerated facial features have the potential of making facial nonmanual signals easier to read. Children find cartoon-like avatars appealing; however adults prefer a more realism in avatar appearance (Kipp, Heloir, & Nguyen, 2011).

Fig. 58 Avatars from ProDeaf, Hand Talk (Hand Talk, 2019)

Fig. 59 Simax, Star from Huawei, and Adamo-Villani's Jason

Table 2 summarizes of past and current signing avatar projects. Many of the projects have Web sites that document the supporting research. See Table 3. Both research projects and commercial products are included. Commercial products include Vcom3D, Simax, StorySign and HandTalk. Empty cells in the table indicate that the authors were not able to find published information.

A last consideration is motion. No matter how improved the appearance, or how accurate the facial pose, the nature of motion is crucial to the clarity and believability of an avatar's signing. To assess the information content of motion in ASL, Poizner et al. (1981) isolated the motion of a signer by first attaching small lights to a fluent signer's shoulders, elbows, wrists and index fingertips, and then recording the signer while in a darkened room. The recordings showed only points of light moving on a dark background. Deaf signers viewed the videos and were asked to identify the signs. The viewers were highly accurate in identifying the signs in the videos. Malala et al. (2018) have also studied the information content of motion. They used various forms of an avatar – one with the hands obscured by spheres, another with the spinal nonmanual signals removed – to assess whether participants could still identify a picture, based on the avatar's signing.

Poizner's and Malala's studies focused primarily on motion of the manual channel of signed discourse. However, since native Deaf signers focus their attention on the face, it would be well worth studying the effect of facial motion as well as facial appearance.

An informal set of evidence in this regard can be found by following the links in Table 4. The first three links are to movies that were created for people to view. These are not interactive, but fixed. This allows the animators to create movement that flows naturally and believably from sign to sign. The second set of links are to recordings of avatars that are interactive. To create these utterances, the supporting application grabs position and motion information for individual words and stitches them together to create a sentence that the user wants to see. The motions in these recordings are not as natural as in the movies.

6. Conclusions and future directions

There has been a great deal of progress in depicting facial nonmanual signals in the past twenty years, but as of yet, there are still unsolved problems impeding researchers from achieving the goal of depicting of facial nonmanual signals that are clear, correct, and easy-to-read.

Improving avatar motion is key. There is need for the fine details that make an avatar believable as a living being, and there is also a need for more insights into larger prosodic units as well. This will require a continuation of collecting and analyzing corpora of signed languages, particularly of the face. Researchers have noted that more precision in measurements of data may lead to further insights into the structure of sign languages (Morrissey & Way, 2013). Acquiring finer detail in corpora will be helpful to both linguists and to the computer scientists developing avatar technology. An additional challenge is understanding linguistic intent of an utterance. Is the motion a result of a single linguistic event or multiple events? What levels of prosody are participating?

There is a need to delve into an even finer level of measurement, that of biomechanical details. Although they do not rise to the level of a linguistic event, the more of these details that avatars can portray when producing utterances, the more believable and readable the utterance will be.

Supporting the coherent and effective leveraging of this data will require research into developing additional sign language representations integrating all levels of detail. For effective avatar animation, having information about all levels of events, from sub-linguistic to prosodic structure will allow avatar developers to create more smoothly flowing animation, approaching that of animated movies. To improve the animation of facial nonmanual signals requires data beyond the pose of the desired nonmanual signal. It also requires the timing of its onset, its intensity and its duration. These may or may not be temporally aligned with lexical entities.

Another challenge is to improve the depiction of prosodic processes. Having access to prosodic data will enable the possibility of creating animations where clausal boundaries are easily distinguished from transitions between lexical items.

Because the goal of signing avatars is to improve Deaf access, it is essential that Deaf communities be involved in avatar development. More involvement with Deaf communities will result in better avatars. It does not matter what hearing developers think; avatar quality can only be determined by the people who are the leading experts in sign languages. Future research directions should take their cue from them.

Table 2 Avatar projects and products

citation	project / product	avatar name	purpose / application	linguistic re-presentation	geometric re-presentation	motion	video or web3D?	Visual / text editing?
Adamo-Villani & Wilbur 2015	Teaching Avatars	Jason	STEM education for Deaf students	ASL-pro notation				both
Bouzid & Jemni, 2014	tunisigner		teaching Sign Writing	Sign Writing	bones	keyframe	web 3D	text
Efthimiou, et al., 2019	sis-builder	UEA avatar	Deaf Ed, Web accessibility, linguistic research	add'l GSL NM HamNoSys	masked blend shapes	keyframe	web 3D	both
Elliott R. et al., 2000	ViSiCAST	UEA avatar	Providing resources for Deaf accessible interfaces	HamNoSys	blend shapes	keyframe mocap	web 3D	text
Gibet et al., 2011	SignCom		corpus creation / improve deaf accessibility to media	hybrid	blend shapes	mocap	video	both
Hand Talk 2020	HandTalk	Hugo	improve deaf access to web sites, automatic translation				web 3D	
Hanke, Popescu & Schmaling, 2003	eSIGN	UEA avatar	Providing resources to create Deaf accessible interfaces	HamNoSys	blend shapes	keyframe	web 3D	both
Heloir & Kipp, 2009	Embodied Agents Behavior Realizer	EMBR	gesture research evaluating signing avatars	EMBRscript	blend shapes	keyframe mocap	web 3D	text
Huawai 2020	StorySign	Star	Deaf education				video	
Matthes et al., 2012	Dicta-Sign	UEA avatar	Developing resources for Deaf accessible interfaces	HamNoSys	masked blend shapes	keyframe	web 3D	both
McDonald et al 2017	Sign Language Avatar at DePaul	Paula	improve Deaf access; promote linguistic research	hybrid	bones	keyframe	video	both
Moemedi & Connan, 2010			English to South African Sign Language (SASL)	Sign Writing	bones	keyframe	web 3D	text
Pauser 2019	Simax	Simax avatar	automatic translation accessible museums				video	
Sims & Silverglate, 2002	Vcom3D	Signing Avatar	Signing Science, Sign 4 Me, Illustrated Dictionary		blend shapes		web 3D	both

Table 3 Avatar projects and products, continued

citation	project Website
Adamo-Villani & Wilbur 2015	http://hpcg.purdue.edu/idealab/
Bouzid & Jemni, 2014	http://tunisigner.com/

link broken; use

http://hpcg.purdue.edu/idealab/

The text may appear identical, but the associated link is compromised.

The link is currently "p//tunisigner.com"

It needs to be "http://tunisigner.com"

Efthimiou, et al., 2019	https://sign.ilsp.gr/sisbuilder/
Elliott R. et al., 2000	http://www.visicast.cmp.uea.ac.uk/Visicast_index.html
Gibet et al., 2011	https://hal.inria.fr/hal-00914661v1
Hand Talk 2020	https://handtalk.me/en
Hanke, Popescu & Schmaling, 2003	http://www.visicast.cmp.uea.ac.uk/eS
Heloir & Kipp, 2009	http://embots.dfki.de/EMBR [1]
Huawai 2020	https://consumer.huawei.com/en/campaign/storysign
Matthes et al., 2012	https://www.sign-lang.uni-hamburg.de/dicta-sign/portal/
McDonald et al 2017	http://asl.cs.depaul.edu
Moemedi & Connan, 2010	
Pauser 2019	https://simax.media
Sims & Silverglate, 2002	http://vcom3d.com

Broken link. Currently "ps//sign.ilsp.gr/sisbuilder" Should be "https://sign.ilsp.gr/sisbuilder/"

http://www.visicast.cmp.uea.ac.uk/Visicast_index.html

https://hal.inria.fr/hal-00914661v1

https://handtalk.me/en

http://www.visicast.cmp.uea.ac.uk/eSIGN/

delete

text on page correct; link broken

text on page correct; link broken

link missing

text on page correct; link broken

Table 4 Samples of avatar singing online

signing

Animated movies of signing	**Title**	**year**	**URL**	**notes**
	World Federation of the Deaf 2007	2007	https://www.youtube.com/watch?v=wW2KBXrPEdM	captioned
	The Forest -- A Story in ASL in HD	2008	https://www.youtube.com/watch?v=oUclQ10BsH8	
	ASL "The forest" (original from youtube user Vcom3D)	2008	https://www.youtube.com/watch?v=80L2Xc0K8Jg	captioned
	My Three Animals	2015	https://vimeopro.com/mocaplab/mocaplab-and-sign-language-avatars/video/122622712	transcript
Recordings of dynamic synthesis	**Title**	**Year**	**URL**	**notes**
	Signing Science Project	2004	https://www.w3.org/WAI/RD/2004/06/sims-mov.htm	
	Say It, Sign It	2007	http://mqtt.org/projects/sisi	voice over
	SignCom Signing Avatar - Story 1	2010	https://www.youtube.com/watch?v=SKEkQIuNpxM	
	ASL Animation with Prosody	2010	http://hpcg.purdue.edu/idealab/web/Jason_W_Prosody_captionsQT.mov	transcript
	A Coffee Story	2012	http://asl.cs.depaul.edu/demo.html	captioned
	Multilingual Speech to Sign Language Translator	2016	https://www.youtube.com/watch?v=fEvvrLpTb0E	captioned

Pls. reformat URLs to be live links/

1 Via the Internet Archive https://archive.org/web/

7. Bibliography

Abel, R. (Director). (1984). *Brilliance* [Motion Picture]. Retrieved March 29, 2020, from https://www.youtube.com/watch?v=7OGlLKzqHHk

Adamo-Villani, N., & Anasingaraju, S. (2017). Holographic signing avatars for deaf education. In *E-Learning, E-Education, and Online Training* (pp. 54–61). Springer.

Adamo-Villani, N., & Wilbur, R. B. (2015). Asl-pro: American sign language animation with prosodic elements. *International Conference on Universal Access in Human-Computer Interaction*, (pp. 307–318).

Adamo-Villani, N., Doublestein, J., & Martin, Z. (2004). The MathSigner: an interactive learning tool for American sign language. *Proceedings. Eighth International Conference on Information Visualisation, 2004. IV 2004.*, (pp. 713–716).

Ahmad, N., Barberis, D., Garazzino, N., Prinetto, P., Shoaib, U., & Tiotto, G. (2012). A virtual character based italian sign language dictionary. *Universal Learning Design, Linz 2012*, 31.

Ahn, J., Gobron, S., Thalmann, D., & Boulic, R. (2013). Asymmetric facial expressions: revealing richer emotions for embodied conversational agents. *Computer Animation and Virtual Worlds, 24*, 539–551.

Alexander, O., Rogers, M., Lambeth, W., Chiang, M., & Debevec, P. (2009). Creating a photoreal digital actor: The digital emily project. *2009 Conference for Visual Media Production*, (pp. 176–187).

Andrei, S., Osborne, L., & Smith, Z. (2013). Designing an American Sign Language avatar for learning computer science concepts for deaf or hard-of-hearing students and deaf interpreters. *Journal of Educational Multimedia and Hypermedia, 22*, 229–242.

Autodesk. (2019). *Things.* Retrieved from Community Creations: https://www.tinkercad.com/things

Azuola, F., Badler, N. I., Ho, P.-H., Huh, S., Kokkevis, E., & Ting, B.-J. (1997). Efforts in Preparation for Jack Validation.

Bahan, B. (1997). Non-manual realization of agreement in American Sign Language. *Non-manual realization of agreement in American Sign Language.*

Baker, C. (1976). Eye-Openers in ASL.

Baker, C. (1978). Focusing on the nonmanual components of American Sign Language. *Understanding language through sign language research.*

Baker, C., & Cokely, D. (1980). American sign language. *A Teacher's Resource Text on Grammar and Culture. Silver Spring, MD: TJ Publ.*

Baker-Shenk, C. (1983). A microanalysis of the nonmanual components of questions in American Sign Language.

Baker-Shenk, C. (1985). The facial behavior of deaf signers: Evidence of a complex language. *American Annals of the Deaf, 130*, 297−304.

Baraki, E. (2010, May 17). *How to draw faces: female.* Retrieved Apr 11, 2020, from Sharenoesis: http://sharenoesis.com/article/draw-face/84

Barberis, D., Garazzino, N., Prinetto, P., & Tiotto, G. (2011). Improving accessibility for deaf people: an editor for computer assisted translation through virtual avatars. *The proceedings of the 13th international ACM SIGACCESS conference on computers and accessibility*, (pp. 253−254).

Bartle, R. A. (2004). *Designing virtual worlds.* New Riders.

Behr, J., Eschler, P., Jung, Y., & Zöllner, M. (2009). X3DOM: a DOM-based HTML5/X3D integration model. *Proceedings of the 14th international conference on 3D web technology*, (pp. 127−135).

Bellugi, U., & Fischer, S. (1972). A comparison of sign language and spoken language. *Cognition, 1*, 173−200.

Benitez-Quiroz, C. F., Gökgöz, K., Wilbur, R. B., & Martinez, A. M. (2014). Discriminant features and temporal structure of nonmanuals in American Sign Language. *PloS one, 9.*

Bergeron, P., Lachapelle, P., Langlois, D., & Robidoux, P. (Directors). (1985). *Tony de Peltrie* [Motion Picture]. Retrieved Mar 30, 2020, from https://www.youtube.com/watch?v=munTr4vmxYE

Bouzid, Y., & Jemni, M. (2013). An Avatar based approach for automatically interpreting a sign language notation. *2013 IEEE 13th International Conference on Advanced Learning Technologies*, (pp. 92−94).

Bouzid, Y., & Jemni, M. (2014). TuniSigner: A virtual interpreter to learn sign writing. *2014 IEEE 14th International Conference on Advanced Learning Technologies*, (pp. 601−605).

Braem, P. B. (2001). Functions of the Mouthing Component in the Signing of Deaf Early and Late Learners of Swiss German Sign Language,". *Foreign Vocabulary in Sign Languages: A Cross-*

Linguistic Investigation of Word Formation, D. Brentari, Ed. Mahwah, NJ: Erlbaum, 1–47.
Braffort, A., & Chételat-Pelé, E. (2011). Analysis and description of blinking in French sign language for automatic generation. *International Gesture Workshop*, (pp. 173–182).
Braffort, A., Bolot, L., & Segouat, J. (2011). Virtual signer coarticulation in Octopus, a Sign Language generation platform. *GW 2011: The 9th International Gesture Workshop.*
Braffort, A., Sansonnet, J.-P., Martin, J.-C., & Verrecchia, C. (2010). Diva, une architecture pour le support des agents gestuels interactifs sur internet. *Technique et Science Informatiques, 29*, 777–806.
Brugman, H., Russel, A., & Nijmegen, X. (2004). Annotating Multi-media/Multi-modal Resources with ELAN. *LREC.*
Brumm, M., Johnson, R., Hanke, T., Grigat, R.-R., & Wolfe, R. (2018). Use of avatar technology for automatic mouth gesture recognition. *SignNonmanuals 2.*
Brutzman, D., & Daly, L. (2010). *X3D: extensible 3D graphics for Web authors.* Elsevier.
Burtnyk, N., & Wein, M. (1976). Interactive skeleton techniques for enhancing motion dynamics in key frame animation. *Communications of the ACM, 19*, 564–569.
Caftanatov, O. (2018). An Anthropomorphic Avatar-based Approach for Virtual Tutoring Software. *Conference on Mathematical Foundations of Informatics*, (pp. 3–16).
Cameron, J. (Producer), Cameron, J. (Writer), & Cameron, J. (Director). (1989). *The Abyss* [Motion Picture]. USA: Twentieth Century Fox. Retrieved Mar 30, 2020, from https://computeranimationhistory-cgi.jimdofree.com/the-abyss-1989/
Campbell, S. R. (2013). Mathematical modeling and virtual environments. In *Modeling Students' Mathematical Modeling Competencies* (pp. 583–593). Springer.
Cassell, J., Vilhjálmsson, H. H., & Bickmore, T. (2004). Beat: the behavior expression animation toolkit. In *Life-Like Characters* (pp. 163–185). Springer.
Catmull, E. (1972). A system for computer generated movies. *Proceedings of the ACM annual conference-Volume 1*, (pp. 422–431).

Catmull, E., & Parke, F. (Directors). (1972). *A Computer Animated Hand* [Motion Picture]. Retrieved March 29, 2020, from https://www.youtube.com/watch?v=fAhyBfLFyNA

Chiriac, I. A., Stoicu-Tivadar, L., & Podoleanu, E. (2015). Comparing video and avatar technology for a health education application for deaf people. *MIE*, (pp. 516–520).

Costa, A. R., Dimuro, G. P., & others. (2001). Supporting deaf sign languages in written form on the web. *10th World Wide Web Conference, Hong Kong, Home Page of Web and Society Track, http://www10. org/cdrom/posters/frame. html. Processing", Written Form on the Web.*

Cox, S., Lincoln, M., Tryggvason, J., Nakisa, M., Wells, M., Tutt, M., & Abbott, S. (2003). The development and evaluation of a speech-to-sign translation system to assist transactions. *International Journal of Human-Computer Interaction, 16*, 141–161.

Craft, B., Hinkle, D., Sedgwick, E., Alkoby, K., Davidson, M. J., Carter, R., . . . others. (2000). An approach to modeling facial expressions used in American Sign Language. *DePaul CTI Research Conference.*

Crasborn, O. A., Van Der Kooij, E., Waters, D., Woll, B., & Mesch, J. (2008). Frequency distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Language & Linguistics, 11*, 45–67.

De Martino, J. M., Silva, I. R., Bolognini, C. Z., Costa, P. D., Kumada, K. M., Coradine, L. C., . . . others. (2017). Signing avatars: making education more inclusive. *Universal Access in the Information Society, 16*, 793–808.

Dimou, A.-L., Goulas, T., & Efthimiou, E. (2011). Grammar/Prosody modelling in greek sign language: towards the definition of built-in sign synthesis rules. *International Gesture Workshop*, (pp. 183–193).

Diwakar, S., & Basu, A. (2008). A multilingual multimedia Indian sign language dictionary tool. *IJCNLP 2008*, 57.

Ebling, S. (2016). *Automatic translation from German to synthesized Swiss German Sign Language.* Ph.D. dissertation, University of Zurich.

Ebling, S., & Glauert, J. (2013). Exploiting the full potential of JASigning to build an avatar signing train announcements.

Efthimiou, E., Fontinea, S.-E., Hanke, T., Glauert, J., Bowden, R., Braffort, A., . . . Goudenove, F. (2010). Dicta-sign–sign language recognition, generation and modelling: a research effort with applications in deaf communication. *Proceedings of the 4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies*, (pp. 80–83).

Efthimiou, E., Fotinea, S.-E., Goulas, T., Kakoulidis, P., Dimou, A.-L., & Vacalopoulou, A. (2016). A complete environment for deaf learner support in the context of mainstream education. *Proceedings of the Conference Universal Learning Design*, *5*, pp. 35–44.

Efthimiou, E., Fotinea, S.-E., Goulas, T., Vacalopoulou, A., Vasilaki, K., & Dimou, A.-L. (2019). Sign Language Technologies and the Critical Role of SL Resources in View of Future Internet Accessibility Services. *Technologies, 7*, 18.

Efthimiou, E., Fotinea, S.-E., Vogler, C., Hanke, T., Glauert, J., Bowden, R., . . . Segouat, J. (2009). Sign language recognition, generation, and modelling: a research effort with applications in deaf communication. *International Conference on Universal Access in Human-Computer Interaction*, (pp. 21–30).

Ekman, R. (1997). *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS).* Oxford University Press, USA.

ElGhoul, O., & Jemni, M. (2011). WebSign: A system to make and interpret signs using 3D Avatars. *Proceedings of the Second International Workshop on Sign Language Translation and Avatar Technology (SLTAT), Dundee, UK*, *23.*

Elhadj, Y. O., Zemirli, Z., & Al-faraj, B. (2012). Towards a unified 3D animated dictionary for Saudi sign language. *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, (pp. 910–917).

Elliott, E. A., & Jacobs, A. M. (2013). Facial expressions, emotions, and sign languages. *Frontiers in psychology, 4*, 115.

Elliott, R., Bueno, J., Kennaway, R., & Glauert, J. (2010). Towards the integrationn of synthetic sl annimation with avatars into corpus annotation tools. *4th Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Language Technologies, Valletta, Malta*, (p. 29).

Elliott, R., Glauert, J. R., & Kennaway, J. R. (2004). A framework for non-manual gestures in a synthetic signing system. *2nd*

Cambridge workshop on universal access and assistive technology, CWUAAT, (pp. 127–136).

Elliott, R., Glauert, J. R., Kennaway, J. R., & Marshall, I. (2000). The development of language processing support for the ViSiCAST project. *Proceedings of the fourth international ACM conference on Assistive technologies*, (pp. 101–108).

Elliott, R., Glauert, J. R., Kennaway, J. R., Marshall, I., & Safar, E. (2008). Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. *Universal Access in the Information Society, 6*, 375–391.

Emanuel, S. (Producer), Weitz, C. (Writer), & Edwards, G. (Director). (2016). *Rogue One: A Star Wars Story* [Motion Picture]. USA: Walt Disney Studios Motion Pictures. Retrieved Mar 30, 2020, from https://youtu.be/HcXTPFgaa6E

EQ4ALL . (2019, Jul 18). *Project Train4ALL.* Retrieved from youtube.com: https://www.youtube.com/watch?v=Vqqg8Nr208g&feature=youtu.be

Feige, K. (Producer), Wright, E. (Writer), & Reed, P. (Director). (2015). *Ant Man* [Motion Picture]. USA: Walt Disney Studios Motion Pictures. Retrieved Mar 30, 2020, from https://computeranimationhistory-cgi.jimdofree.com/ant-man-2015/

Fergusson, M. B. (1940). Quo Sensu Credis Et Ore?: A Study of Facial Expression in Greek and Latin Literature. *Greece & Rome, 9*, 102–116.

Ferreira, M. A., & García, L. S. (2018). Requirements for Avatar in Virtual Environment of Support Learning in the Literacy of Deaf People in Portuguese Mediated by LIBRAS. *Proceedings of the 17th Brazilian Symposium on Human Factors in Computing Systems*, (pp. 1–6).

Filhol, M., & Hadjadj, M. N. (2016). Juxtaposition as a form feature; syntax captured and explained rather than assumed and modelled. *Language resources and evaluation conference (LREC), Representation and processing of Sign Languages.* Portorož, Slovenia.

Filhol, M., & Hadjadj, M. N. (2016). Juxtaposition as a form feature; syntax captured and explained rather than assumed and modelled. *Language resources and evaluation conference (LREC),* *Representation and processing of Sign Languages.* Portorož, Slovenia.

Filhol, M., & McDonald, J. (2018). Extending the AZee-Paula shortcuts to enable natural proform synthesis. *8th Workshop on the Representation and Processing of Sign Languages:*, 45-52.

Foster, H. (2019, September 12). *Non-Manual Markers in ASL / North Carolina Division of Services for the Deaf and Hard of Hearing .* Retrieved from You Tube: https://www.youtube.com/watch?v=8HIc0IRe-dE&feature=youtu.be

Furst, J., Alkoby, K., Lancaster, G., McDonald, J., & Wolfe, R. (2002). Making Airport Security Accessible to the Deaf. *Proceedings of the Fifth IASTED International Conference on Computer Graphics and Imaging* (pp. 38-43). Kaua'i, HI: IASTED.

Garchery, S., Boulic, R., & Capin, T. (2004). *Standards for Virtual Human Animation: H-ANIM and MPEG4.* Tech. rep., John Wiley.

Gibet, S., Courty, N., Duarte, K., & Naour, T. L. (2011). The SignCom system for data-driven animation of interactive virtual signers: Methodology and Evaluation. *ACM Transactions on Interactive Intelligent Systems (TiiS), 1*, 1–23.

Gibet, S., Lefebvre-Albaret, F., Hamon, L., Brun, R., & Turki, A. (2016). Interactive editing in French Sign Language dedicated to virtual signers: requirements and challenges. *Universal Access in the Information Society, 15*, 525–539.

Glauert, J., & Elliott, R. (2011). Extending the SiGML Notation–a Progress Report. *Second International Workshop on Sign Language Translation and Avatar Technology (SLTAT), 23.*

Good, C. (2016, Sep 11). *When CGI artists say it takes "X" amount of time to render a frame, are they talking literally or figuratively?* Retrieved from Quora: https://www.quora.com/When-CGI-artists-say-it-takes-X-amount-of-time-to-render-a-frame-are-they-talking-literally-or-figuratively

Goulas, T., Fotinea, S.-E., Efthimiou, E., & Pissaris, M. (2010). SiS-Builder: a sign synthesis support tool. *LREC*, (pp. 102–105).

Grieve-Smith, A. B. (1999). English to American Sign Language machine translation of weather reports. *Proceedings of the Second High Desert Student Conference in Linguistics (HDSL2), Albuquerque, NM*, (pp. 23–30).

Grossman, R. B., & Kegl, J. (2006). To capture a face: A novel technique for the analysis and quantification of facial expressions in American Sign Language. *Sign Language Studies, 6*, 273–305.

Grubb, W. (2009). Facial animation rig for Delgo. *Tutorial Notes, 12.*

Gutiérrez, M., Vexo, F., & Thalmann, D. (2002). A mpeg-4 virtual human animation engine for interactive web based applications. *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*, (pp. 554–559).

Hand Talk. (2019). *Make your company website accessible with our LIBRAS translation plugin.* Retrieved from Hand Talk: http://www.handtalk.me/en

Hand Talk. (2019, Jan 29). *ProDeaf changed address (In Portuguese).* Retrieved from ProDeaf: http://www.prodeaf.net/

Hanke, T. (2002). iLex-A tool for Sign Language Lexicography and Corpus Analysis. *LREC.*

Hanke, T. (2004). HamNoSys -- Representing sign language data in language resources and language processing contexts. *Fourth International Conference on Language Resources and Evaluation (LREC 2004). Representation and Processing of Sign Languages Workshop* (pp. 1-6). Paris: European Language Resources Association.

Hanke, T., & Storz, J. (2008). iLex – A Database Tool for Integrating Sign Language Corpus Linguistics and Sign Language Lexicography. *Workshop on the Representation and Processing of Sign Language, at the Sixth International Conference on Language Resources and Evaluation (LREC'08)* (pp. 64-67). Marrakech, Morocco: European Language Resources Association (ELRA).

Hanke, T., Popescu, H., & Schmaling, C. (2003). eSIGN–HPSG-assisted Sign Language Composition. *Gesture Workshop.*

Hayward, K., Adamo-Villani, N., & Lestina, J. (2010). A computer animation system for creating deaf-accessible math and science curriculum materials.

Heloir, A., & Kipp, M. (2009). Embr–a realtime animation engine for interactive embodied agents. *International Workshop on Intelligent Virtual Agents*, (pp. 393–404).

Heloir, A., & Kipp, M. (2010). Real-Time Animation of Interactive Agents: Specification and Realization. *Applied Artificial Intelligence, 24*, 510–529.

Hoffmann-Dilloway, E. (2011). Writing the smile: Language ideologies in, and through, sign language scripts. *Language & Communication, 31*, 345–355.

Holmes, A. (2015, Jul 23). *How Ant-Man actually de-aged Michael Douglas.* Retrieved Mar 30, 2020, from CinemaBlend: https://www.cinemablend.com/new/How-Ant-Man-Actually-De-Aged-Michael-Douglas-73107.html

Huawei. (2020). *StorySign: Helping Deaf children to learn to read.* Retrieved from StorySign -- Huawei Global: https://consumer.huawei.com/en/campaign/storysign/

Huenerfauth, M. (2009). A linguistically motivated model for speed and pausing in animations of american sign language. *ACM Transactions on Accessible Computing (TACCESS), 2*, 1–31.

Huenerfauth, M., & Kacorri, H. (2015). Augmenting EMBR virtual human animation system with MPEG-4 controls for producing ASL facial expressions. *International Symposium on Sign Language Translation and Avatar Technology, 3.*

Jaballah, K., & Jemni, M. (2013). A Review on 3D signing avatars: Benefits, uses and challenges. *International Journal of Multimedia Data Engineering and Management (IJMDEM), 4*, 21–45.

Jamrozik, D. G., Davidson, M. J., McDonald, J. C., & Wolfe, R. (2010). Teaching students to decipher fingerspelling through context: a new pedagogical approach. *Proceedings of the 17th National Convention Conference of Interpreter Trainers, San Antonio, TX*, (pp. 35–47).

Jemni, M., & Elghoul, O. (2008). Using ICT to teach sign language. *2008 Eighth IEEE International Conference on Advanced Learning Technologies*, (pp. 995–996).

Jemni, M., & Elghoul, O. (2008). Using ICT to teach sign language. *2008 Eighth IEEE International Conference on Advanced Learning Technologies*, (pp. 995–996).

Jennings, V., Elliott, R., Kennaway, R., & Glauert, J. (2010). Requirements for a signing avatar. *Proc. Workshop on Corpora and Sign Language Technologies (CSLT), LREC*, (pp. 33–136).

Kacorri, H., & Huenerfauth, M. (2016). Continuous profile models in ASL syntactic facial expression synthesis. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, (pp. 2084–2093).

Kacorri, H., Lu, P., & Huenerfauth, M. (2013). Evaluating facial expressions in American Sign Language animations for accessible online information. *International Conference on Universal Access in Human-Computer Interaction*, (pp. 510–519).

Karpouzis, K., Caridakis, G., Fotinea, S.-E., & Efthimiou, E. (2007). Educational resources and implementation of a Greek sign language synthesis architecture. *Computers & Education, 49*, 54–74.

Kato, M. (2008). A study of notation and sign writing systems for the Deaf. *Intercultural Communication Studies, 17*, 97–114.

Kennaway, J., Glauert, R., & Zwitserlood, I. (2007). Providing signed content on the Internet by synthesized animation. *ACM Transactions on Computer-Human Interaction (TOCHI), 14*(3), 15-1 - 15-26.

Kennaway, R. (2001). Synthetic animation of deaf signing gestures. *International Gesture Workshop*, (pp. 146–157).

Kipp, M., Heloir, A., & Nguyen, Q. (2011). Sign language avatars: Animation and comprehensibility. *International Workshop on Intelligent Virtual Agents*, (pp. 113–126).

Kipp, M., Heloir, A., Schröder, M., & Gebhard, P. (2010). Realizing multimodal behavior. *International Conference on Intelligent Virtual Agents*, (pp. 57–63).

Kipp, M., Nguyen, Q., Heloir, A., & Matthes, S. (2011). Assessing the deaf user perspective on sign language avatars. *The proceedings of the 13th international ACM SIGACCESS conference on Computers and accessibility*, (pp. 107–114).

Klassen, A. (2017, Jul 11). *"Jurassic Park" Turns 20: Film memories with Ariana Richards.* Retrieved Mar 30, 2020, from Daily Beast: https://www.thedailybeast.com/jurassic-park-turns-20-film-memories-with-ariana-richards?ref=scroll

Koch, J.-P. (2016, July 5). *Python – Tool for automating multilingual lip sync.* Retrieved from Jan-Philipp Koch: https://janphilippkoch.wordpress.com/2016/07/05/tool-for-automating-multilingual-lip-sync-using-python-and-maya-api/

Koller, O., Zargaran, S., Ney, H., & Bowden, R. (2018). Deep sign: enabling robust statistical continuous sign language recognition via hybrid CNN-HMMs. *International Journal of Computer Vision, 126*, 1311–1325.

Kouremenos, D., Fotinea, S.-E., Efthimiou, E., & Ntalianis, K. (2010). A prototype Greek text to Greek Sign Language conversion system. *Behaviour & Information Technology, 29*, 467–481.

Krňoul, Z., Kanis, J., Železný, M., & Müller, L. (2007). Czech text-to-sign speech synthesizer. *International Workshop on Machine Learning for Multimodal Interaction*, (pp. 180–191).

Kuroda, T., Sato, K., & Chihara, K. (1998). S-TEL: An avatar based sign language telecommunication system. *International journal of virtual reality, 3*, 20–26.

Lasseter, J. (Producer), & Lasseter, J. (Director). (1986). *Luxo, Jr.* [Motion Picture]. Retrieved Mar 29, 2020, from https://www.youtube.com/watch?v=zmhZm5FRV4s

Lasseter, J. (1987). Principles of traditional animation applied to 3D computer animation. *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, (pp. 35–44).

Lasseter, J. (Director). (1988). *Tin Toy* [Motion Picture]. Retrieved Mar 29 , 2020, from https://www.youtube.com/watch?v=ffIZSAZRzDA

Lext, J., & Akenine-Möller, T. (2001). Towards rapid reconstruction for animated ray tracing. *Eurographics 2001–Short Presentations*, 311-318.

Liddell, S. K. (1980). *American sign language syntax* (Vol. 52). Mouton De Gruyter.

Lombardo, V., Battaglino, C., Damiano, R., & Nunnari, F. (2011). An avatar-based interface for the italian sign language. *2011 International Conference on Complex, Intelligent, and Software Intensive Systems*, (pp. 589–594).

Lucas, C., & Valli, C. (1992). *Language Contact in the American Deaf Community*. San, Diego, CA: Academic Press.

Malala, V. D., Prigent, E., Braffort, A., & Berret, B. (2018). Which Picture? A Methodology for the Evaluation of Sign Language Animation Understandability.

Marmor, D. (2011, Oct 22). *Facial Rigging Blend Shape*. Retrieved Mar 30, 2020, from CG Feedback: http://www.cgfeedback.com/cgfeedback/showthread.php?t=2119

Marshall, I., & Safar, E. (2005). Grammar development for sign language avatar-based synthesis. *Proceedings HCII*, (pp. 1–10).

Martin, G. C. (2018, May 4). *Preston Blair phoneme series.* Retrieved from Gary C. Martin: CG imagery development: http://www.garycmartin.com/mouth_shapes.html

Matsumoto, D., & Hwang, H. S. (2011). Reading facial expressions of emotion. *Psychological Science Agenda, 25.*

Matthes, S., Hanke, T., Regen, A., Storz, J., Worseck, S., Efthimiou, E., . . . Safar, E. (2012). Dicta-Sign–building a multilingual sign language corpus. *Proc. of the 5th Workshop on the Representation and Processing of Sign Languages: Interactions Between Corpus and Lexicon (LREC), European Language Resources Association.*

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. *International Conference on Universal Access in Human-Computer Interaction*, (pp. 41–54).

McIntire, M. L., & Reilly, J. S. (1988). Nonmanual Behaviors in L 1 & L 2 Learners of American Sign Language. *Sign Language Studies, 61*, 351–375.

Moemedi, K., & Connan, J. (2010). Rendering an animated avatar from SignWriting notation. *Proceedings of the Southern Africa Telecommunication Networks and Applications Conference.*

Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33-35.

Morrissey, S., & Way, A. (2013). Manual labour: tackling machine translation for sign languages. *Machine Translation, 27*, 25–64.

Nadolske, M. A., & Rosenstock, R. (2007). Occurrence of mouthings in American Sign Language: A preliminary study. *Visible variation: Comparative studies on sign language structure*, 35–61.

Naqvi, S. (2007). End-user Involvement in Assistive Technology Design for the Deaf-are Artificial Forms of Sign Language Meeting the Needs of the Target Audience? *CVHI.*

Noll, A. (Director). (1965). *Computer Ballet* [Motion Picture]. Retrieved March 25, 2020, from https://computeranimationhistory-cgi.jimdofree.com/computer-ballet-1965/

Noll, A. M. (1965). *Computer-generated three-dimensional movies.* Bell Telephone Laboratories.

Oh, Y.-J., Park, K.-H., & Bien, Z. (2007). Body motion editor for sign language avatar. *2007 International Conference on Control, Automation and Systems*, (pp. 1752–1757).

Oleva, S. (2018, Jan). *Tin Toy (1988).* Retrieved Mar 30, 2020, from History of Computer Animation (CGI): https://computeranimationhistory-cgi.jimdofree.com/tin-toy-1988/

Orvalho, V., Bastos, P., Parke, F. I., Oliveira, B., & Alvarez, X. (2012). A Facial Rigging Survey. *Eurographics (STARs)*, (pp. 183–204).

Pandzic, I. S., & Forchheimer, R. (2002). The Origins of the MPEG-4 Facial Animation Standard. *MPEG-4 Facial Animation: The Standard, Implementation And Applications*, 3–13.

Papadogiorgaki, M., Grammalidis, N., Makris, L., Sarris, N., & Strintzis, M. G. (2004). VSigns–a virtual sign synthesis web tool. *Workshop on Information and Knowledge Management for Integrated Media Communication*, (pp. 25–31).

Pauser, S. (2019, Mar 19). *Prototypentest SiMAX im Rahmen des Innovationsschecks.* Retrieved from Equalizent: https://www.equalizent.com/fileadmin/user_upload/News/2019_04_Avatar_Projektbericht.pdf

Pfau, R., & Quer, J. (2010). Nonmanuals: Their prosodic and grammatical roles. *Sign languages*, 381–402.

Piekartz, H. v., & Mohr, G. (2014). Reduction of head and face pain by challenging lateralization and basic emotions: a proposal for future assessment and rehabilitation strategies. *Journal of Manual & Manipulative Therapy, 22*, 24–35.

Pluralsight Creative. (2013, Jun 24). *Maya Tutorial: The Anatomy of an Expression for Facial Animation in Maya.* Retrieved Apr 18, 2020, from YouTube: https://www.youtube.com/watch?v=dkE_CQwSdhk

Poizner, H., Bellugi, U., & Lutes-Driscoll, V. (1981). Perception of American sign language in dynamic point-light displays. *Journal of experimental psychology: Human perception and performance, 7*, 430.

Prillwitz, S., & für Deutsche Gebärdensprache und Kommunikation Gehörloser, H. Z. (1989). *Hamnosys: Version 2.0; hamburg notation system for sign languages; an introductory guide.* Signum-Verlag.

Reilly, I., & Anderson, D. (2002). in ASL. *Directions in sign language acquisition, 2*, 159.

Reilly, J. S., McIntire, M., & Bellugi, U. (1990). The acquisition of conditionals in American Sign Language: Grammaticized facial expressions. *Applied Psycholinguistics, 11*, 369–392.

Royce, B. (2019, Aug 21). *The Matrix is getting a fourth movie – so can we have The Matrix Online back now please?* Retrieved Apr 16, 2020, from Massively Overpowered: https://massivelyop.com/2019/08/21/the-matrix-is-getting-a-fourth-movie-so-can-we-have-the-matrix-online-back-now-please/

Sallandre, M. (2007). Simultaneity in French sign language discourse. *Amsterdam Studies in the Theory and History of Linguistic Science Series 4, 281*, 103.

Sanders, A.-L. (2019, Nov 17). *Flash Frame-By-Frame Animation: 8-Frame Basic Walk Cycle.* Retrieved Mar 30, 2020, from Lifewire: https://www.lifewire.com/flash-basic-walk-cycle-140629

San-Segundo, R., Barra, R., Córdoba, R., d'Haro, L. F., Fernández, F., Ferreiros, J., . . . Pardo, J. M. (2008). Speech to sign language translation system for Spanish. *Speech Communication, 50*, 1009–1020.

Sansonnet, J.-P., Braffort, A., Segouat, J., & Verrecchia, C. (2009). SL Teacher: a Framework for teaching LSF on the Web. *ICTA*, (pp. 189–196).

Schleifer, J., Scaduto-Mendola, R., Canetti, Y., & Piretti, M. (2002). Character setup from rig mechanics to skin deformations: A practical approach. *SIGGRAPH Course Notes, 2.*

Schnepp, J. (2006). Towards Computerized Synthesis of Nonmanual Signals in American Sign Language. *DePaul CTI Research Symposium (CTIRS) & Midwest Software Engineering Conference (MSEC), Chicago, Illinois.*

Schnepp, J., Wolfe, R., McDonald, J., & Toro, J. (2013). Generating Co-occurring Facial Nonmanual Signals in Synthesized American Sign Language. *Eight International Conference on Computer Graphics Theory and Applications*, (pp. 407-416). Barcelona.

Segouat, J., Braffort, A., Bolot, L., Choisier, A., Filhol, M., & Verrecchia, C. (2008). Building 3D French Sign Language lexicon. *Workshop Programme*, (p. 174).

Segouat, J., Braffort, A., Bolot, L., Choisier, A., Filhol, M., & Verrecchia, C. (2008). Building 3D French Sign Language lexicon. *Workshop Programme*, (p. 174).

Shoaib, U., Ahmad, N., Prinetto, P., & Tiotto, G. (2014). Integrating multiwordnet with Italian sign language lexical resources. *Expert systems with applications, 41*, 2300–2308.

Shohieb, S. M. (2019). A Gamified e-Learning Framework for Teaching Mathematics to Arab Deaf Students:(Supporting an Acting Arabic Sign Language Avatar). *Ubiquitous Learning: An International Journal, 12*, 55–70.

Shumaker, C. (2016, Feb 2). *NMS Facial Expression.* Retrieved Apr 15, 2020, from YouTube: https://www.youtube.com/watch?v=NbbNwVwdfGg

Silver, J. (Producer), Wachowski, L., & Wachowski, L. (Directors). (2003). *The Matrix Revolutions* [Motion Picture]. USA: Warner Bros. Retrieved Mar 30, 2020, from https://youtu.be/Sh906GSszWk

Sims, E., & Silverglate, D. (2002). Interactive 3D characters for web-based learning and accessibility. *ACM SIGGRAPH 2002 conference abstracts and applications*, (pp. 314–314).

Sinha, R. (2014, Jun 5). *Call of Duty: Advanced Warfare Uses Avatar 2's Facial Animation System.* Retrieved Mar 30, 2020, from Gaming Bolt: https://gamingbolt.com/call-of-duty-advanced-warfare-uses-avatar-2s-facial-animation-system

Spielberg, S. (Director). (1993). *Jurassic Park* [Motion Picture]. Retrieved Mar 30, 2020, from https://youtu.be/BPUEqvfQuHY

Stokoe, W. C. (1960). Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf (Studies in Linguistics, Occasional Papers 8). *Buffalo, NY: University of Buffalo.*

Sutton, V. (2013, Jul 1). *SignWriting Workshop in LaJolla, San Diego California.* Retrieved from SignWriting.org: https://www.signwriting.org/archive/docs9/sw0824_SignWriting_Workshop_Flyer_LaJolla_July2013.pdf

Toni. (2019, Feb 15). *Jump Force Special Editions Revealed and Avatar Customization Screenshots.* Retrieved Apr 11, 2020, from Video Games Blogger: https://www.videogamesblogger.com/2018/11/21/jump-force-special-editions-revealed-and-avatar-customization-screenshots.htm

Toro, J. A., McDonald, J. C., & Wolfe, R. (2014). Fostering better deaf/hearing communication through a novel mobile app for fingerspelling. *International Conference on Computers for Handicapped Persons*, (pp. 559–564).

Van Der Hulst, H., & Channon, R. (2010). Notation systems. In D. Brentari (Ed.), *Sign Languages* (pp. 151–172). Cambridge University Press.

Van Zijl, L., & Combrink, A. (2006). The South African sign language machine translation project: issues on non-manual sign generation. *Proceedings of the 2006 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, (pp. 127–134).

Verlinden, M., Tijsseling, C., & Frowein, H. (2001). A Signing Avatar on the WWW. *International Gesture Workshop*, (pp. 169–172).

Verlinden, M., Zwitserlood, I., & Frowein, H. (2005). Multimedia with animated sign language for deaf learners. *EdMedia+ Innovate Learning*, (pp. 4759–4764).

Vesel, J. (2005, Mar 15). *Signing Math & Science.* Retrieved from TERC: https://www.terc.edu/projects/signing-math-science/

Vesel, J., & Robillard, T. (2013). Teaching mathematics vocabulary with an interactive signing math dictionary. *Journal of Research on Technology in Education, 45*, 361–389.

Vlasic, D., Brand, M., Pfister, H., & Popovic, J. (2006). Face transfer with multilinear models. In *ACM SIGGRAPH 2006 Courses* (pp. 246–433).

Vogler, C., & Goldenstein, S. (2005). Analysis of facial expressions in american sign language. *Proc, of the 3rd Int. Conf. on Universal Access in Human-Computer Interaction, Springer.*

Wailes, P. (2014, Oct 23). *Visual Maths: A Brief History of CGI.* Retrieved Apr 16, 2020, from Internet Archive: https://web.archive.org/web/20160314051025/https://builtvisible.com/visual-maths-brief-history-cgi/

Waller, B. M., Cray Jr, J. J., & Burrows, A. M. (2008). Selection for universal facial emotion. *Emotion, 8*, 435.

Wang, N., Gao, X., Tao, D., Yang, H., & Li, X. (2018, Jan). Wang, N., Gao, X., Tao, D., Yang, H., & Li, X. (2018). Facial feature point detection: A comprehensive survey. *Neurocomputing, 275*, 50-65.

WatchMojo.com. (2019, Jun 22). Top 10 Deepfake Videos. USA. Retrieved Mar 30, 2020, from https://www.youtube.com/watch?v=-QvIX3cY4lc

Weast, T. P. (2008). Questions in American Sign Language: A quantitative analysis of raised and lowered eyebrows.

Web 3D Consortium. (1999). *Home*. Retrieved from Web 3D: open Standards for Real-time 3D communication: https://www.web3d.org/

Wilbur, R. (1994). Eyeblinks & ASL phrase structure. *Sign Language Studies, 84*, 221–240.

Wilbur, R. B. (2009). Effects of varying rate of signing on ASL manual signs and nonmanual markers. *Language and speech, 52*, 245–285.

Wolfe, R., Efthimiou, E., Glauert, J., Hanke, T., McDonald, J., & Schnepp, J. (Eds.). (2016). Special Issue: recent advances in sign language translation and avatar technology. *Universal Access in the Information Society.*

Xu, Y., Feng, A. W., Marsella, S., & Shapiro, A. (2013). A practical and configurable lip sync method for games. In *Proceedings of Motion on Games* (pp. 131–140).

Zemeckis, R. (Director). (1992). *Death Becomes Her* [Motion Picture]. Retrieved Mar 30, 2020, from https://youtu.be/p2hZia7wU-Q

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N., & Palmer, M. (2000). A machine translation system from English to American Sign Language. *Conference of the Association for Machine Translation in the Americas*, (pp. 54–67).