Evaluation of Animated Swiss German Sign Language Fingerspelling Sequences and Signs

Sarah Ebling[1], Sarah Johnson[2], Rosalee Wolfe[2], Robyn Moncrief[2], John McDonald[2], Souad Baowidan[2], Tobias Haug[3], Sandra Sidler-Miserez[3], and Katja Tissi[3]

[1]University of Zurich, Zurich, Switzerland
[2]DePaul University, Chicago, IL, USA
[3]University of Applied Sciences of Special Needs Education (HfH), Zurich, Switzerland
ebling@cl.uzh.ch,sarahej101@gmail.com,{wolfe,jmcdonald}@cs.depaul.edu,
{rkelley5,sbaowida}@mail.depaul.edu,{tobias.haug,katja.tissi}@hfh.ch,
sandysidler@gmail.com

Abstract. This paper reports on work in animating Swiss German Sign Language (DSGS) fingerspelling sequences and signs as well as on the results of a study evaluating the acceptance of the animations. The animated fingerspelling sequences form part of a fingerspelling learning tool for DSGS, while the animated signs are to be used in a study exploring the potential of sign language avatars in sign language assessment. To evaluate the DSGS fingerspelling sequences and signs, we conducted a focus group study with seven early learners of DSGS. We identified the following aspects of the animations as requiring improvement: non-manual features (in particular, facial expressions and head and shoulder movements), (fluidity of) manual movements, and hand positions of fingerspelling signs.

Keywords: Sign language animation, animation evaluation, focus group

1 Introduction

Sign language animation, the process of creating a signing avatar, is a young field of research, looking back on about 20 years of existence. In contrast to videos of human signers, sign language animations are capable of providing an anonymous representation of a signer. This minimizes the likelihood of legal implications arising from, e.g., display on the web. Moreover, the content of a sign language animation can typically be modified more easily than that of a self-contained video. Using sign language animation also bears the possibility of tailoring the avatar's appearance (gender, level of formality, etc.) and speed of signing to a user's needs [7].

Sign language animation can be used as part of tools for learning *finger alphabets*, i.e., communication systems associated with sign languages in which

dedicated signs are used for each letter of a spoken language[1] word. Figure 1 shows the finger alphabet of Swiss German Sign Language (*Deutschschweizerische Gebärdensprache*, DSGS). Note that it features separate signs for -Ä-, -Ö-, and -Ü- as well as for -CH- and -SCH-. Traditional fingerspelling learning tools display one still image corresponding to the prototypical hand configuration for each letter of a sequence, thereby merely visualizing the holds (static postures) of that sequence. In contrast, sign language animation is capable of accounting for all of the salient information inherent in fingerspelling, namely both holds and transitions (movements).
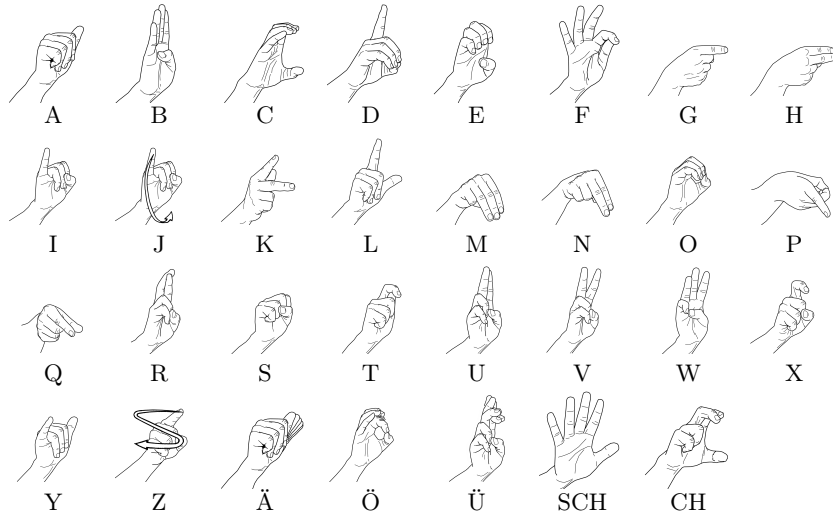
Fig. 1. Finger alphabet of DSGS [3]

We hypothesize that sign language avatars additionally have the potential to increase motivation and interest in young learners of sign language, thus evoking the *Persona effect* [15] that has been observed in pedagogical agents for children in spoken languages. Our aim is to conduct a study in which the items of a Receptive Skills Test (RST) for DSGS [10] are signed by an avatar instead of a human. The DSGS RST is completed by Deaf children between ages four and eleven. The test assesses morphological constructions of DSGS such as spatial verb morphology, negation, number, distribution, and verb agreement through 46 items. In its current form, an item consists of a video of a human signer performing a DSGS sequence, such as BÄR KLEIN ('BEAR SMALL'), APFEL VIELE ('APPLE MANY'), or BUB SCHAUEN-oben ('BOY LOOK-upward'). Test takers are then asked to pick among three or four images the correct one,

---

[1] *Spoken language* refers to a language that is not signed, represented as speech or text.

i.e., the image that best matches the content previously signed. Figure 2 shows the options given for the sequence APFEL VIELE, where B is the targeted response.
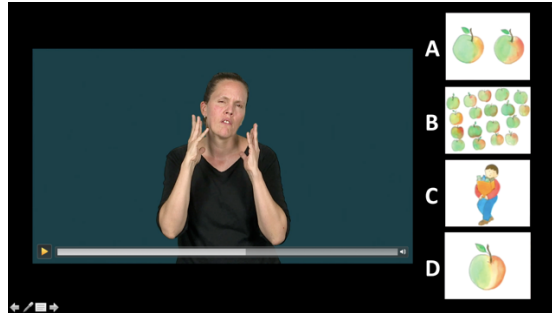
Fig. 2. Item APFEL VIELE ('APPLE MANY') in the DSGS Receptive Skills Test [10]

The first step in our study designed to gauge the potential of sign language avatars in sign language assessment consists of creating animations of the test items.[2] This paper reports on the results of a study evaluating the acceptance of 1) a subset of these animations and 2) DSGS fingerspelling sequences generated through the DSGS fingerspelling learning tool described above. The study was a focus group conducted with seven early learners[3] of DSGS and a Deaf[4] moderator.

The remainder of this paper is structured as follows: Section 2 discusses previous work on sign language animation (Section 2.1) and evaluation of sign language animation (Section 2.2). Section 3 presents work on animation of DSGS fingerspelling sequences (Section 3.1) and signs (Section 3.2). Section 4 presents the setup and the results of the focus group study conducted to evaluate the DSGS sign and fingerspelling animations. Finally, Section 5 offers a conclusion and an outlook on future work.

---

[2] The 46 test items contain 91 unique signs, both lexicalized and productive.

[3] This term refers to Deaf persons who learned DSGS in a family situation from birth or as pre-schoolers/in their very first primary school years. Limiting the subject pool to native signers is generally not an option for DSGS due to the small population upon which one could draw.

[4] It is a widely recognized convention to use the upper-cased word *Deaf* for describing members of the linguistic community of sign language users and the lower-cased word *deaf* for describing the audiological state of a hearing loss [17].

2   Previous Work

2.1   Sign Language Animation

Sign language animations are typically created through one of three approaches: animation by hand (traditional animation), motion capturing, or fully synthesized (procedural) animation. Animation by hand consists of manually modelling and posing an avatar character in a purpose-built tool or commercially/freely available software such as Maya, 3ds Max, or Blender. This procedure is highly labor-intensive but generally yields very good results. A signing avatar may also be animated based on information obtained from motion capturing, which involves recording a human's signing. While the quality of sign language animations obtained through motion capturing tends to be high, major drawbacks of this approach are the long calibration time and the extensive postprocessing required.

Both with hand-crafted animation and with animation from motion capturing, the inventory of available signing comprises precisely the sign forms previously created and their combinations. The sublexical structure of the signs is usually not accessible at runtime. This is different for the fully synthesized animation approach: Here, animations are created from a gesture/form notation, which means at execution time there is access to the sublexical structure of signs at whatever level of detail the underlying notation system offers. In case of, e.g., the Hamburg Notation System for Sign Languages (HamNoSys) [9,19], the place of articulation of a sign and other parameters can be adjusted on the fly to take account of coarticulation effects. The fact that fully synthesized animation allows for signs to be modified in context *ad hoc* renders it the most flexible of the three approaches to sign language animation. At the same time, this approach typically results in the lowest quality, as controlling the appearance of all possible sign forms that may be produced from a given notation inventory is virtually impossible [5].

The avatar that is at the core of the work described in this paper, Paula, relies on both hand animation and procedural animation. Paula has been used to develop a fingerspelling learning tool for American Sign Language (ASL) [22,24,28], which served as the basis for the DSGS fingerspelling learning tool described in Section 3.1. While individual signs in ASL can be used to represent whole words, fingerspelling allows signers to spell out names, proper nouns, acronyms, and other words for which there is no explicit sign [18]. In a survey of hearing students learning ASL, participants cited the understanding of fingerspelling as the most challenging aspect of learning the language [23]. While a student may learn the alphabet early in their coursework, spelling a word in practice involves not just each individual letter, but also the movement of the transitions between them [1], giving the entire word a unique shape [8]. This was the main motivation in designing the ASL fingerspelling software mentioned above, which includes not just still frames of each letter, but also animated transitions to more accurately replicate native fingerspelling. However, there is no way for an animator to individually recreate every possible word, and real-

time procedural generation of the transitions carries a high computational cost. In response to these problems, unique transitions between every possible two-letter combination in the alphabet were hand-animated. This makes it possible to create any arbitrary word with greatly reduced computational requirements, natural movement, and no awkward penetrations between the fingers during the transitions from one shape to the next [26].

Similar to the fingerspelling software, an ASL sentence generator [25] takes individual, hand-animated signs as a motion base, and procedurally transitions between them to form unique sentences using any combination of pre-animated signs. The sentence generator has the additional feature of allowing procedural incorporation of modifiers in order to scale the range of the movement, change its speed, and, in the future, convey a wide variety of emotions. The use of human-generated animation reduces the robotic movements that seem mechanical and awkward to Deaf people, similar to how computer-generated voices sound stilted and unnatural to hearing people. The procedural automation allows for cheap and quick generation of arbitrary sentences and phrases without the exorbitant time and labor costs of hand animation.

2.2   Sign Language Animation Evaluation

No automatic procedure exists for assessing the quality of signing avatars. Sign language animation evaluation studies so far have been carried out in the form of user studies. Here, a distinction is typically made between two concepts: the degree to which a user understands the content of an animation (*comprehension*) and the degree to which he or she accepts it (*acceptance*) [11]. While these two concepts cannot be taken to be independent (most importantly, comprehension is likely to affect acceptance), distinguishing between them makes sense in light of the method used to assess each concept: Comprehension is typically assessed through objective comprehension tasks, while acceptance is commonly assessed via subjective participant judgments. Several studies assessing the comprehension of signing avatars have been carried out [11,12,16,20,14,21]. [13] conducted what is to date the most comprehensive acceptance study; other acceptance studies include those of [4,16,20]. The study introduced in this paper (cf. Section 4) represents an acceptance study as well.
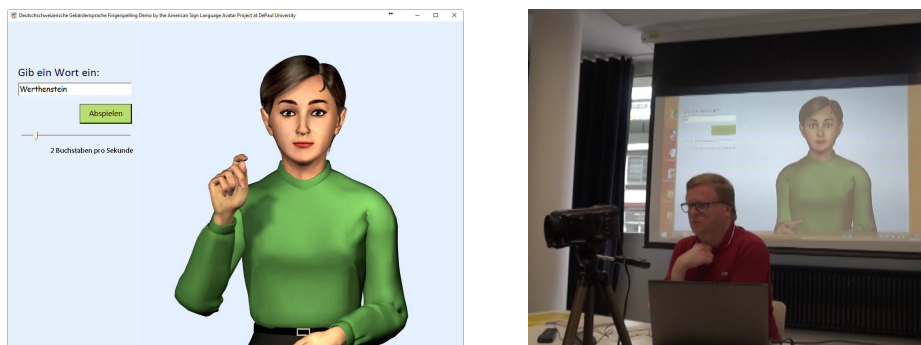
3   DSGS Animation

3.1   Animation of DSGS Fingerspelling Sequences

Departing from work on an ASL fingerspelling learning tool (cf. Section 2.1), development of a DSGS fingerspelling learning tool has recently begun [6]. Just like with ASL, synthesizing the DSGS finger alphabet consisted of producing hand postures (handshapes with orientations) for each letter of the alphabet (as shown in Figure 1) and transitions for each pair of letters. Recall that the DSGS finger alphabet contains signs for -Ä-, -Ö-, -Ü-, -CH-, and -SCH-, which

are not present in the ASL finger alphabet. In addition, four handshapes, -F-, -G-, -P-, and -T-, are distinctly different from ASL. Further, the five letters -C-, -M-, -N-, -O-, and -Q- have a similar handshape in DSGS, but required smaller modifications, such as a different orientation or small adjustments in the fingers. Hence, the DSGS finger alphabet features 14 out of 30 hand postures that needed modification from the ASL finger alphabet.

We conducted an online study to assess the comprehensibility of the resulting animated DSGS fingerspelling sequences among Deaf and hearing participants; details of this study are given in [6]. Participants saw 22 names of places in Switzerland fingerspelled by either a human or the Paula signing avatar and were asked to type the letters of the word in a text box. The resulting comprehension rate of the signing avatar was highly satisfactory at 90.06%. In the general comments section, one participant encouraged the introduction of speed controls for the signing avatar.

While the participants of the study reported on in [6] were shown isolated videos of animated DSGS fingerspelling sequences, in the meantime, a Desktop interface for Windows similar to the one available for ASL[5] has been developed (cf. Figure 3). This interface was demonstrated to the participants of the focus group study described in Section 4. Among other functionality, the interface offers the possibility of adjusting the speed of fingerspelling, a feature implemented in response to the previous user study.

Fig. 3. Fingerspeller interface and presentation in the focus group

3.2   Animation of DSGS Signs

While the ASL and DSGS finger alphabets are similar, there are some key differences between these two languages. For one, DSGS heavily relies on mouthing,[6] i.e., making a mouth movement as if to pronounce a spoken language word but
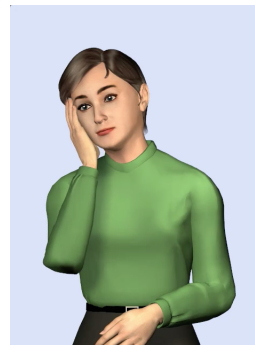
---

[5] http://fingerspellingtutor.com/
[6] According to [2], 80-90% of signs in DSGS are accompanied by a mouthing.

with no vocalization, along with the manual sign for full comprehension. This process seems to be less dominant in ASL. For an initial expansion into DSGS, we animated ten signs using a transcriber software [27] that allows for animating lexical signs by hand in a linguistically informed way: APFEL ('APPLE'), AUTO ('CAR'), BALL ('BALL'), BÄR ('BEAR'), BETT ('BED'), BLEISTIFT ('PEN-CIL'), BUB ('BOY'), DA ('THERE'), ESSEN ('EAT'), and FRAU ('WOMAN'). The ten signs were from the DSGS RST described in Section 1. Figure 4 shows a screenshot of the animation of BETT.

For this preliminary study, we did not include the mouthing motions, as we were more focused on portraying the mechanics of the hands and body as accurately as possible for review by our focus group. We understand the importance of this part of the language, and fully intend to include it in future animations. Other non-manual features such as head and upper-body movements (cf. Figure 4) were included.

Fig. 4. Animation of DSGS sign BETT ('BED')

Additionally, the transcriber software described above was expanded to include a wider variety of handshapes based on HamNoSys. This allows users more flexibility in animating a variety of different sign languages, making the overall animation work-flow faster and more efficient.

4   Focus Group Study

We conducted a focus group study to evaluate the acceptance of the animated DSGS fingerspelling sequences and signs described in Sections 3.1 and 3.2. Seven early learners of DSGS between ages 32 and 55 participated in the study, all but one of whom were certified DSGS instructors working for the Swiss Deaf Association. The study took place on the premises of the Swiss Deaf Association in Zurich and was moderated by a Deaf DSGS user not affiliated with our research. The study took 1.5 hours.

The focus group consisted of three activities: Firstly, participants were shown three examples of avatars producing continuous signing and asked to evaluate them (cf. Figure 5). The avatars presented were Paula signing content in ASL, Mira by Braam Jordaan with many features of sign language poetry,[7] and an avatar produced by MocapLab in collaboration with Gallaudet University sign-ing the ASL nursery rhyme "My Three Animals".[8] The avatars had been selected to represent different possible use cases of avatars; this was to stimulate a dis-cussion among the participants as to what additional fields of application of sign language avatars could be. Participants were then presented with ten animated DSGS signs from the DSGS RST (cf. Section 3.2) and asked for their feedback.

---

[7] http://braamjordaan.com/

[8] http://www.mocaplab.com/projects/gallaudet-university/

Following this, the moderator demonstrated the DSGS fingerspelling learning tool (cf. Section 3.1) and again solicited feedback.

Fig. 5. Three avatars shown: Paula, Mira, and MocapLab/Gallaudet

When evaluating the three different avatars (Activity 1), participants stressed the importance of facial expressions, of which they stated Mira had a lot (one participant even deemed it too much), the MocapLab/Gallaudet avatar a bit less, and Paula too little. Mira's expressiveness was the reason why this avatar was not envisioned in a public information setting (e.g., for conveying train or air travel announcements). The participants agreed that Paula would be most suitable for such purposes.

With regard to the ten DSGS signs (Activity 2), participants pointed out the lack of mouthings. Another aspect that came up for the majority of signs shown was facial expression, of which the participants requested there should be more. Additionally, they wished to see more movement of the head, shoulders, and upper body for two of the signs (APFEL, AUTO). Regarding manual activity, mention was made of some movements being executed too abruptly: For example, with BLEISTIFT, the movement back to neutral position at the end of the sign was taken to be too instantaneous, and the initial movement of the hand to the ear in FRAU was judged as being too fast. The handshapes and/or hand positions of some signs were also deemed as needing improvement.

For the fingerspelling learning tool (Activity 3), the participants remarked that many of the handshapes were correct but that some hand positions (e.g., of -P- and -D-) were not. They also commented on the absence of *glides*, i.e., single executions of a letter combined with a horizontal movement to represent double letters (as opposed to two successive executions of the letter). Similarly, they noted that while a single sign for -SCH- existed, the sign was not used in the fingerspelling sequences but instead replaced with S-C-H. Further feedback targeted the height at which some fingerspelling signs were executed. For example, the participants remarked that the signing location of -M-, -N-, and -Q- was too low.

5   Conclusion and Outlook

This paper has reported on work in animating DSGS fingerspelling sequences and signs as well as on the results of a study evaluating the acceptance of the an-

imations. We have described ongoing work in developing a DSGS fingerspelling learning tool and including sign language animations in sign language assessment. As a result of the focus group study we conducted, we identified the following aspects of the animations as being in need of improvement: non-manual features (in particular, facial expressions as well as head and shoulder movements), (fluidity of) manual movements, and hand positions of fingerspelling signs.

Our future work will focus on improving these aspects. In addition, we will implement routines that replace S-C-H and C-H with -SCH- and -CH-, respectively, where appropriate and incorporate glides for double letters in the fingerspeller interface.

Acknowledgments. We are grateful to the Swiss Deaf Association for hosting the focus group.

References

1. Akamatsu, C.: Fingerspelling formulae: A word is more or less than the sum of its letters. In: Stokoe, W., Volterra, V. (eds.) Proceedings of Sign Language Research '83. pp. 126–132. Linstok Press, Silver Spring, MD (1985)
2. Boyes Braem, P.: Functions of the mouthing component in the signing of Deaf early and late learners of Swiss German Sign Language. In: Brentari, D. (ed.) Foreign Vocabulary in Sign Languages: A Cross-Linguistic Investigation of Word Formation, pp. 1–47. Erlbaum, Mahwah, NJ (2001)
3. Boyes Braem, P.: A multimedia bilingual database for the lexicon of Swiss German Sign Language. Sign Language & Linguistics 4(1/2), 133–143 (2001)
4. Ebling, S.: Evaluating a Swiss German Sign Language avatar among the Deaf community. In: Proceedings of the 3rd International Symposium on Sign Language Translation and Avatar Technology (SLTAT). Chicago, IL (2013), retrieved from `http://www.zora.uzh.ch/85717/1/CAMERA_READY_sltat2013_submission_14.pdf` (last accessed November 20, 2015)
5. Ebling, S.: Automatic Translation from German to Synthesized Swiss German Sign Language. Ph.D. thesis, University of Zurich (2016)
6. Ebling, S., Wolfe, R., Schnepp, J., Baowidan, S., McDonald, J., Moncrief, R., Sidler-Miserez, S., Tissi, K.: Synthesizing the finger alphabet of Swiss German Sign Language and evaluating the comprehensibility of the resulting animations. In: Proceedings of the 6th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT). Dresden, Germany (2015)
7. Glauert, J.: Animating sign language for Deaf people (2013), lecture held at the University of Zurich, October 9, 2013 (unpublished)
8. Groode, J.: Fingerspelling: Expressive and Receptive Fluency. Dawn Sign Press, San Diego (1992), videotape.
9. Hanke, T.: HamNoSys – representing sign language data in language resources and language processing contexts. In: Proceedings of the 1st LREC Workshop on Representation and Processing of Sign Languages. pp. 1–6. Lisbon, Portugal (2004)
10. Haug, T., Perollaz, R.: Verständnistest Deutschschweizer Gebärdensprache (2015)

11. Huenerfauth, M., Zhao, L., Gu, E., Allbeck, J.: Evaluating American Sign Language generation through the participation of native ASL signers. In: Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS). pp. 211–218. Tempe, AZ (2007)

12. Kipp, M., Heloir, A., Nguyen, Q.: Sign language avatars: Animation and comprehensibility. In: Proceedings of the 11th International Conference on Intelligent Virtual Agents (IVA). pp. 113–126. Reykjavík, Iceland (2011)

13. Kipp, M., Nguyen, Q., Heloir, A., Matthes, S.: Assessing the Deaf user perspective on sign language avatars. In: Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS). pp. 107–114. Dundee, Scotland (2011)

14. Lefebvre-Albaret, F.: DictaSign Deliverable D7.4: Prototype evaluation synthesis. Tech. rep., DictaSign project (2011)

15. Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A., Bhogal, R.S.: The persona effect: Affective impact of animated pedagogical agents. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI). pp. 359–366. Atlanta, GA (1997)

16. McDonald, J., Wolfe, R., Moncrief, R., Baowidan, S.: A computational model of role shift to support the synthesis of signed language, presented at the 12th Theoretical Issues in Sign Language Research (TISLR), Melbourne, Australia, January 4-7 (2016)

17. Morgan, G., Woll, B.: The development of complex sentences in British Sign Language. In: Morgan, G., Woll, B. (eds.) Directions in Sign Language Acquisition: Trends in Language Acquisition Research, pp. 255–276. John Benjamins, Amsterdam, Netherlands (2002)

18. Padden, C.: The acquisition of fingerspelling by deaf children. In: Siple, P., Fischer, S. (eds.) Theoretical Issues in Sign Language Research. pp. 191–210. University of Chicago, Chicago (1991)

19. Prillwitz, S., Leven, R., Zienert, H., Hanke, T., Henning, J.: HamNoSys: Version 2.0: An Introductory Guide. Signum, Hamburg, Germany (1989)

20. Schnepp, J., Wolfe, R., McDonald, J., Toro, J.: Generating Co-occurring Facial Nonmanual Signals in Synthesized American Sign Language. In: Proceedings of the Eighth International Conference on Computer Graphics Theory and Applications (GRAPP), Barcelona, Spain, February 21-24. pp. 407–416 (2013)

21. Smith, R., Nolan, B.: Emotional Facial Expressions in Synthesised Sign Language Avatars: A Manual Evaluation. Universal Access in the Information Society pp. 1–10 (2015)

22. Toro, J.A., McDonald, J., Wolfe, R.: Fostering better Deaf/hearing communication through a novel mobile app for fingerspelling. In: Proceedings of the 14th International Conference on Computers Helping People with Special Needs (ICCHP). pp. 559–564. Paris, France (2014)

23. Wilcox, S.: The Phonetics of Fingerspelling. John Benjamins, Amsterdam, Netherlands (1992)

24. Wolfe, R., Alba, N., Billups, S., Davidson, M.J., Dwyer, C., Gorman Jamrozik, D., Smallwood, L., Alkoby, K., Carhart, L., Hinkle, D., Hitt, A., Kirchman, B., Lancaster, G., McDonald, J., Semler, L., Schnepp, J., Shiver, B., Suh, A., Young, J.: An Improved Tool for Practicing Fingerspelling Recognition. In: Proceedings of the International Conference on Technology and Persons with Disabilities. pp. 17–22. Los Angeles, CA (2006)

25. Wolfe, R., Cook, P., McDonald, J.C., Schnepp, J.: Linguistics as structure in computer animation: Toward a more effective synthesis of brow motion in American Sign Language. Sign Language & Linguistics 14(1), 179–199 (2013)
26. Wolfe, R., McDonald, J., Davidson, M., Frank, C.: Using an Animation-based Technology to Support Reading Curricula for Deaf Elementary Schoolchildren. In: Proceedings of the 22nd Annual International Technology and Persons with Disabilities Conference. Los Angeles, CA (2007)
27. Wolfe, R., McDonald, J., Schnepp, J.: An Avatar to Depict Sign Language: Building from Reusable Hand Animation, presented at the International Workshop on Sign Language Translation and Avatar Technology (SLTAT), Federal Ministry of Labour and Social Affairs, Berlin, Germany, January 10-11 (2011)
28. Wolfe, R., McDonald, J., Toro, J., Baowidan, S., Moncrief, R., Schnepp, J.: Promoting Better Deaf/Hearing Communication Through an Improved Interaction Design for Fingerspelling Practice. In: Universal Access in Human-Computer Interaction. Access to Today's Technologies, pp. 495–505. Springer (2015)