Understanding ASL Learners' Preferences for a Sign Language Recording and Automatic Feedback System to Support Self-Study

Saad Hassan
Computing and Information Science
Rochester Institute of Technology
Rochester, NY, USA
sh2513@rit.edu

Sooyeon Lee
School of Information
Rochester Institute of Technology
Rochester, NY, USA
slics@rit.edu

Dimitris Metaxas
Computer Science
Rutgers University
Piscataway, NJ, USA
dnm@rutgers.edu

Carol Neidle
Linguistics
Boston University
Boston, MA, USA
carol@bu.edu

Matt Huenerfauth
School of Information
Rochester Institute of Technology
Rochester, NY, USA
matt.huenerfauth@rit.edu

ABSTRACT

Advancements in AI will soon enable tools for providing automatic feedback to American Sign Language (ASL) learners on some aspects of their signing, but there is a need to understand their preferences for submitting videos and receiving feedback. Ten participants in our study were asked to record a few sentences in ASL using software we designed, and we provided manually curated feedback on one sentence in a manner that simulates the output of a future automatic feedback system. Participants responded to interview questions and a questionnaire eliciting their impressions of the prototype. Our initial findings provide guidance to future designers of automatic feedback systems for ASL learners.

CCS CONCEPTS

• **Human-centered computing** → *Interaction design process and methods*; *Accessibility design and evaluation methods*; • **Applied computing** → *Computer-managed instruction.*

KEYWORDS

Sign languages, Education, Feedback, American Sign Language, Language learning, Automatic feedback, Interface design

ACM Reference Format:
Saad Hassan, Sooyeon Lee, Dimitris Metaxas, Carol Neidle, and Matt Huenerfauth. 2022. Understanding ASL Learners' Preferences for a Sign Language Recording and Automatic Feedback System to Support Self-Study. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '22), October 23–26, 2022, Athens, Greece.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3517428.3550367

1 INTRODUCTION AND RELATED WORK

American Sign Language (ASL) is used by about half a million people in the U.S. as a primary form of communication [21]. ASL classes have one of the fastest growing language-learning enrollments; it is now the 3rd most studied non-native language in the U.S. [1, 10], with 200,000 students [9] as of 2018. Individuals motivated to learn ASL include future ASL interpreters, hearing parents and teachers of Deaf and Hard of hearing (DHH) children, teachers in bilingual ASL-English programs, and other professionals [25, 28].

One common method for providing feedback to students in ASL classes is for an instructor to watch and provide feedback on each student's signing production during class. For homework or out-of-class assignments, instructors may watch a video recording of the student and provide written feedback. In some educational settings, instructors also employ specialized software that enables time-synchronized feedback to users, e.g. GoReact [1] and TerpTube [16]. Providing feedback imposes a heavy workload on instructors, and students are unable to obtain rapid feedback, which may limit their learning [17]. Advancements in machine learning have enabled automatic detection of individual signs or aspects of signing in video, e.g., [15, 30, 31]; a recent system can identify some errors in the production of ASL sentences with moderate accuracy (60%) [26]. Other advancements have given rise to new methods for providing more customized feedback to students, e.g., using visual feedback to point to specific parts of the body where the error occurs [23, 24]. Advances in deep-fake face-swapping technologies, e.g., [22], could be applied to a video of an expert signer, so that the student could see their own face on the body of the expert signer, to help visualize how they would appear performing ASL in a correct manner.

Providing students with automatic feedback during language-learning assignments has been explored in a variety of contexts and languages [3, 6, 14, 29]. However, relatively little work has focused on sign language [5, 18]. There has also been some research on developing ASL games and other engaging interactive learning tools [4, 7, 8, 19, 25, 27], as well as sign language dictionaries [2, 11–13]. However, these technologies are mostly designed to improve receptive, rather than expressive, skills of learners. Given the advances in AI-powered technologies discussed above, there is a need to understand how students would like to engage with automatic feedback tools, especially how they would like to submit videos and view feedback. In this poster paper, we investigate:

RQ1 What are the expectations of ASL learners from software that lets them record a video to get feedback?

ASSETS '22, October 23–26, 2022, Athens, Greece

ACM ISBN 978-1-4503-9258-7/22/10.
https://doi.org/10.1145/3517428.3550367

RQ2 What are the subjective preferences of ASL students with respect to viewing various forms of automatically generated feedback?

2 METHODOLOGY

2.1 Participants and Recruitment

This IRB-approved study was conducted remotely using Zoom video-conferencing because of the COVID-19 pandemic. We reached out to professors teaching introductory ASL courses, who shared an advertisement by email with their students, and we also posted on Reddit. Six participants identified as women, and 4 as men. (An option for non-binary and other self-described gender options was provided.) The mean age of our participants was 26.8 years. Eight participants had fewer than 2 years of experience in learning ASL.

2.2 Study Protocol

Participants were informed that the purpose of this study was to understand their experiences in learning ASL or taking ASL classes, recording video assignments, and obtaining feedback. The study consisted of four phases. Three phases occurred during the initial appointment: a *pre-recording interview*; *time for the participant to record videos of their own signing*; and *a post-recording interview*. Finally, a few days after the appointment, *they received a video with feedback and were asked to respond to a post-study questionnaire.*

Participants were first asked about their prior experiences in recording ASL videos and getting feedback. Participants then interacted with a prototype created for this study. The software allowed students to play some sample videos of expert signers performing ASL sentences or short passages consisting of 3-4 sentences. These stimuli videos were from the curriculum of *Signing Naturally*, a popular textbook used for ASL instruction [20]. Participants were able to browse and replay these videos multiple times, and then, when ready, they were asked to make a video recording of themselves mimicking each video; the video played while they were recording in a side-window. The recording-software prototype used the webcam on the participants' local computer to record their ASL signing, and the prototype enabled the students to record their own videos while the videos of the expert signers were playing, so that they could follow along.

A post interview was conducted to get feedback on the recording software. Finally, participants were told to expect to receive an email in a few days with a video of their own signing with feedback added, and they were told that they would need to complete a questionnaire after watching that feedback video. One of our co-authors who is an ASL linguistics expert carefully inspected one video from each participant and prepared feedback videos in three different formats, as shown in Figures 2 and 3.

(1) By pointing to a location on the face or body of the student when displaying text messages on the screen about how the ASL signing may be correct or incorrect.
(2) By zooming in on the face of the ASL student at times when there was an error in the production of a facial expression.
(3) By displaying the result of face-swap software that overlaid the face of the participant onto the face of an expert ASL signer (with the participant's face conveying appropriate facial expression movements based on the expert ASL signer), to illustrate students would appear if they were producing ASL sentences in a correct manner.

We sent a video with feedback to each participant along with the link to an online questionnaire, asking participants about their impression of the feedback provided in the video.

2.3 Data Analysis

To analyze participants' interview and questionnaire responses, we employed both deductive and inductive coding approaches in our qualitative data analysis. Based on our research aims, we aggregated participants' comments about the recording software and the feedback provided separately. We performed a thematic analysis on both of these sets of comments separately. We present our findings in sections 3.1 and 3.2 respectively.

3 RESULTS

3.1 Side-by-side Presentation

Participants' feedback was largely positive regarding their experience of watching a video of an expert signer while recording themselves reproducing the signing.

3.1.1 Less worrying about vocabulary. Participants appreciated not having to worry about the vocabulary and focusing more on the fluency of signing. For example, P1 commented: *I liked being able to imitate it and see exactly what I needed to do next, instead of like, completely remembering the sentence, especially if I didn't know the signs.*

3.1.2 Adjusting signing on the fly. P5 also liked the fact that they could *adjust (their) signs on the fly to more what the signer was doing.* They could also focus more on other aspects of fluent signing. P2 commented: *It's the easiest way to compare myself to the signer. Am I maintaining the right attitude? Am I using the space around me properly? Are my signs too big or too small?*

3.1.3 Pace signing better. Two participants mentioned that watching a video while recording themselves allowed them to pace their signing more like a native signer. To support this further P5 also suggested: *time stamp suggestions, where it's not only giving you the timestamp of where you as the signer differed from the video, but also getting the timestamps of the video where they're doing it correctly. Just like making direct comparisons is as easy as possible.*

3.1.4 Distraction. Three participants expressed concerns about the side-by-side presentation. For example, P3 said that it distracted them and they ended up focusing more on the native signer's video and missing facial expressions: *I was focusing too much on them rather than like trying to look at it, get it down and make sure that I was there. I feel like I was leaving off facial expressions.*

P4 agreed: *I felt like I wasn't signing naturally cause I was looking at it and trying to match what they were saying.*

Five participants wanted to review their own recording immediately, e.g., P7 wanted to: *see myself back after, because I'm generally trying not to pay attention to myself on the screen while I'm signing.*

One participant wanted a tool to provide immediate translations of the stimulus video, saying: *So I think there's not really any translator, what the sign it is, maybe? - P9*

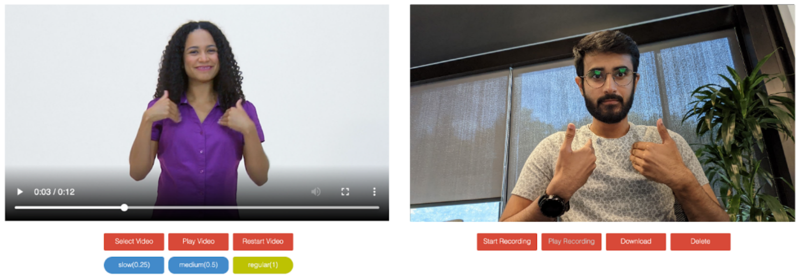
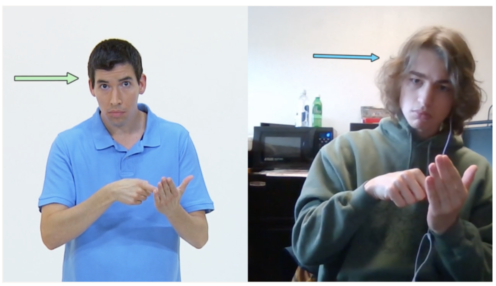


Figure 1: Screenshots of the recorder software with two windows. The window on the left shows the video presented to the participant. The right window is the viewfinder, showing the video from the webcam of the user's computer.



(a) Feedback using an arrow pointing to the error location

(b) Zooming in on face to emphasize the moment at which an error in facial expression occurs

Figure 2: Screenshots of manually curated feedback videos shared with our participants, with an expert signer shown on left and participant (who granted permission for their face to appear in this paper) on the right.

3.2 Feedback

At least six participants liked the side-by-side presentation of the original video and the student's video in the feedback provided, e.g.: *I really liked being able to see the signer and me side-by-side. Sometimes I forget ... to ... raise eyebrows or lean forward, and having that visual with the differences pointed out was really helpful. I feel like just this small activity will make me more aware of what I should be doing when signing. - P1*

Participants pointed out the long-term benefits of using a tool like this: *continued use of this system could help me fine-tune my sign, especially my nonmanual skills. - P8*

Figure 3: Manually curated feedback video with expert signer (left), expert signer with their face-swapped with participant (middle), and participant (right, permission granted for face to appear in this paper) in a side-by-side composite manner.

Participants also discussed how they liked the feedback on facial expressions, especially early in signing education, e.g.: *it is crucial to have this tool implemented, as signing habits can be developed earlier on when the focus is more directed towards picking up signs, understanding the grammar etc. And less on subtle things like [non-manual] markers.*

3.2.1 Text-based Feedback. Participants liked text-based feedback when it was succinct, e.g. *It was short and sweet. The next comment was long. There were many suggestions, and a lot to keep track of while watching a quick sentence. - P1*

Participants also liked the zooming-in aspect: *Showing the side-by-side comparison, especially when zooming in on the error, was helpful. - P8*

3.2.2 Visual Indicators. Overall, participants appreciated the visual indicators used in feedback videos and offered suggestions on how to improve them, e.g. P1 said, *I liked the arrows that pointed out differences. There were a lot of large red arrows for the second segment that felt a little cluttered and clunky. I think a skinnier red arrow might help indicate the location I should be looking at better.*

3.2.3 Face Swap. There were two participants who commented positively on the face-swap technology. P4 said it was an *encouragement for the things that were done correctly.* Others, e.g. P10, liked the idea but said the face-swap technology needs to improve: *I liked* *the idea of it because I feel like it can improve. I disliked the execution of how the face was just planted on.*

4 LIMITATIONS, FUTURE WORK, AND CONCLUSION

There are several limitations of this work that can serve as basis for future studies. The number of participants (10) in this study was relatively small. Future studies, can recruit a larger set of participants. Our feedback videos were manually curated by one of our co-authors. The feedback was also not shared with our participants in real time since we needed time to manually curate the feedback. Once technologies that can detect errors in signing are more readily available, researchers can use them to provide automatically generated feedback to participants. Future research can look at more alternatives for presentation of feedback to learners.

The goal of this research was to capture experiences of different ASL learners and their opinions regarding a system that can provide them with feedback on their signing. Our preliminary findings may inform future designers of such technologies about the requirements of ASL learners for such tools.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grants 1763569, 2040638, 2212303, 2125362.

REFERENCES

[1] Jodie M Ackerman, Ju-Lee A Wolsey, M Diane Clark, et al. 2018. Locations of L2/Ln Sign Language Pedagogy. *Creative Education* 9, 13 (2018), 2037.

[2] Oliver Alonzo, Abraham Glasser, and Matt Huenerfauth. 2019. Effect of Automatic Sign Recognition Performance on the Usability of Video-Based Search Interfaces for Sign Language Dictionaries. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) *(ASSETS '19)*. Association for Computing Machinery, New York, NY, USA, 56–67. https://doi.org/10.1145/3308561.3353791

[3] Pierre Badin, Atef Ben Youssef, Gérard Bailly, Frédéric Elisei, and Thomas Hueber. 2010. Visual articulatory feedback for phonetic correction in second language learning. In *L2SW, Workshop on" Second Language Studies: Acquisition, Learning, Education and Technology*. P1–10.

[4] Dhruva Bansal, Prerna Ravi, Matthew So, Pranay Agrawal, Ishan Chadha, Ganesh Murugappan, and Colby Duke. 2021. *CopyCat: Using Sign Language Recognition to Help Deaf Children Acquire Language Skills*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411763.3451523

[5] Helene Brashear. 2007. Improving the efficacy of automated sign language practice tools. *ACM SIGACCESS Accessibility and Computing* 89 (2007), 11–17.

[6] Marinella Caruso, Nicola Fraschini, and Sabine Kuuse. 2019. Online Tools for Feedback Engagement in Second Language Learning. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)* 9, 1 (2019), 58–78.

[7] Kirsten Ellis, Neil Ray, and Cheryl Howard. 2011. Learning a physical skill via a computer: a case study exploring Australian Sign Language. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*. 98–103.

[8] Julie Fisher, Kirsten Ellis, Louisa Willoughby, and Jan Carlo Barca. 2014. Taking a user centred design approach for designing a system to teach sign language. ACIS.

[9] National Center for Education Statistics (NCES). 2018. Digest of Education Statistics Number and percentage distribution of course enrollments in languages other than English at degree-granting postsecondary institutions, by language and enrollment level: Selected years, 2002 through 2016. https://nces.ed.gov/programs/digest/d18/tables/dt18_311.80.asp

[10] David Goldberg, Dennis Looney, and Natalia Lusin. 2015. Enrollments in Languages Other than English in US. Institutions of Higher Ed., Fall'13.

[11] Saad Hassan, Oliver Alonzo, Abraham Glasser, and Matt Huenerfauth. 2020. Effect of ranking and precision of results on users' satisfaction with search-by-video sign-language dictionaries. In *Sign Language Recognition, Translation and Production (SLRTP) Workshop-Extended Abstracts*, Vol. 4. Computer Vision – ECCV 2020 Workshops, Virtual, 6 pages.

[12] Saad Hassan, Oliver Alonzo, Abraham Glasser, and Matt Huenerfauth. 2021. Effect of Sign-Recognition Performance on the Usability of Sign-Language Dictionary Search. *ACM Trans. Access. Comput.* 14, 4, Article 18 (oct 2021), 33 pages. https://doi.org/10.1145/3470650

[13] Saad Hassan, Akhter Al Amin, Alexis Gordon, Sooyeon Lee, and Matt Huenerfauth. 2022. Design and Evaluation of Hybrid Search for American Sign Language to English Dictionaries: Making the Most of Imperfect Sign Recognition. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 195, 13 pages. https://doi.org/10.1145/3491102.3501986

[14] Saad Hassan, Aiza Hasib, Suleman Shahid, Sana Asif, and Arsalan Khan. 2019. Kahaniyan - Designing for Acquisition of Urdu as a Second Language. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernando Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis Zaphiris (Eds.). Springer International Publishing, Cham, 207–216.

[15] Tobias Haug, Sarah Ebling, Katja Tissi, Sandra Sidler-Miserez, and Penny Boyes Braem. 2022. Development of a Technology-Assisted Assessment for Sign Language Learning. *Intl. Journal of Emerging Technologies in Learning* 17 (2022).

[16] Evan Hibbard, Martin Gerdzhev, Tara Stevens, Daniel Roush, and Deborah Fels. 2020. Getting a Sign in Edgewise: User-Centered Design Considerations in Creating a Signed Language Mentoring Management System. *Sign Language Studies* 20, 2 (2020), 264–300.

[17] Matt Huenerfauth, Elaine Gale, Brian Penly, Sree Pillutla, Mackenzie Willard, and Dhananjai Hariharan. 2017. Evaluation of Language Feedback Methods for Student Videos of American Sign Language. *ACM Trans. Access. Comput.* 10, 1, Article 2 (apr 2017), 30 pages. https://doi.org/10.1145/3046788

[18] Matt Huenerfauth, Elaine Gale, Brian Penly, Mackenzie Willard, and Dhananjai Hariharan. 2015. Comparing Methods of Displaying Language Feedback for Student Videos of American Sign Language. In *Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility* (Lisbon, Portugal) *(ASSETS '15)*. Association for Computing Machinery, New York, NY, USA, 139–146. https://doi.org/10.1145/2700648.2809859

[19] Teerawat Kamnardsiri, Ler-on Hongsit, Pattaraporn Khuwuthyakorn, and Noppon Wongta. 2017. The Effectiveness of the Game-Based Learning System for the Improvement of American Sign Language using Kinect. *Electronic Journal of e-Learning* 15, 4 (2017), pp283–296.

[20] Ella Mae Lentz. 1994. " Signing Naturally": An American Sign Language Curriculum. *Parallel Views: Education and Access for Deaf People in France and the United States* (1994), 131.

[21] Ross Mitchell, Travas Young, Bellamie Bachleda, and Michael Karchmer. 2006. How Many People Use ASL in the United States? Why Estimates Need Updating. *Sign Language Studies* 6 (03 2006). https://doi.org/10.1353/sls.2006.0019

[22] Yuval Nirkin, Yosi Keller, and Tal Hassner. 2019. FSGAN: Subject Agnostic Face Swapping and Reenactment. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*.

[23] Han Duy Phan, Kirsten Ellis, and Alan Dorin. 2018. MIC, an interactive sign language teaching system. In *Proceedings of the 30th Australian Conference on Computer-Human Interaction*. 544–547.

[24] Han Duy Phan, Kirsten Ellis, Alan Dorin, and Patrick Olivier. 2020. Feedback strategies for embodied agents to enhance sign language vocabulary learning. In *Proceedings of the 20th ACM International Conf. on Intelligent Virtual Agents*. 1–8.

[25] Jerry Schnepp, Rosalee Wolfe, Gilbert Brionez, Souad Baowidan, Ronan Johnson, and John McDonald. 2020. Human-Centered Design for a Sign Language Learning Application. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (Corfu, Greece) *(PETRA '20)*. Association for Computing Machinery, New York, NY, USA, Article 60, 5 pages. https://doi.org/10.1145/3389189.3398007

[26] Elahe Vahdani, Longlong Jing, Yingli Tian, and Matt Huenerfauth. 2021. Recognizing American Sign Language Nonmanual Signal Grammar Errors in Continuous Videos. In *2020 25th International Conference on Pattern Recognition (ICPR)*. 1–8. https://doi.org/10.1109/ICPR48806.2021.9413126

[27] Kimberly A. Weaver, Harley Hamilton, Zahoor Zafrulla, Helene Brashear, Thad Starner, Peter Presti, and Amy Bruckman. 2010. Improving the Language Ability of Deaf Signing Children through an Interactive American Sign Language-Based Video Game. In *Proceedings of the 9th International Conference of the Learning Sciences - Volume 2* (Chicago, Illinois) *(ICLS '10)*. International Society of the Learning Sciences, 306–307.

[28] Kimberly A. Weaver and Thad Starner. 2011. We Need to Communicate! Helping Hearing Parents of Deaf Children Learn American Sign Language. In *The Proceedings of the 13th International ACM SIGACCESS Conference on Computers and Accessibility* (Dundee, Scotland, UK) *(ASSETS '11)*. Association for Computing Machinery, New York, NY, USA, 91–98. https://doi.org/10.1145/2049536.2049554

[29] Jin Ha Woo and Heeyoul Choi. 2021. Systematic Review for AI-based Language Learning Tools. *arXiv preprint arXiv:2111.04455* (2021).

[30] Yuancheng Ye, Yingli Tian, Matt Huenerfauth, and Jingya Liu. 2018. Recognizing American Sign Language Gestures From Within Continuous Videos. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) Workshops*.

[31] Chenyang Zhang, Yingli Tian, and Matt Huenerfauth. 2016. Multi-modality American Sign Language recognition. In *2016 IEEE International Conference on Image Processing (ICIP)*. 2881–2885. https://doi.org/10.1109/ICIP.2016.7532886