Evaluation of Language Feedback Methods for Student Videos of American Sign Language

MATT HUENERFAUTH, Rochester Institute of Technology
ELAINE GALE, Hunter College, City University of New York
BRIAN PENLY, SREE PILLUTLA, MACKENZIE WILLARD,
and DHANANJAI HARIHARAN, Rochester Institute of Technology

This research investigates how to best present video-based feedback information to students learning American Sign Language (ASL); these results are relevant not only for the design of a software tool for providing automatic feedback to students but also in the context of how ASL instructors could convey feedback on students' submitted work. It is known that deaf children benefit from early exposure to language, and higher levels of written language literacy have been measured in deaf adults who were raised in homes using ASL. In addition, prior work has established that new parents of deaf children benefit from technologies to support learning ASL. As part of a long-term project to design a tool to automatically analyze a video of a students' signing and provide immediate feedback about fluent and non-fluent aspects of their movements, we conducted a study to compare multiple methods of conveying feedback to ASL students, using videos of their signing. Through two user studies, with a Wizard-of-Oz design, we compared multiple types of feedback in regard to users' subjective judgments of system quality and the degree students' signing improved (as judged by an ASL instructor who analyzed recordings of students' signing before and after they viewed each type of feedback). The initial study revealed that displaying videos to students of their signing, augmented with feedback messages about their errors or correct ASL usage, yielded higher subjective scores and greater signing improvement. Students gave higher subjective scores to a version in which time-synchronized pop-up messages appeared overlaid on the student's video to indicate errors or correct ASL usage. In a subsequent study, we found that providing images of correct ASL face and hand movements when providing feedback yielded even higher subjective evaluation scores from ASL students using the system.

CCS Concepts: • **Human-centered computing → User studies**; **Empirical studies in accessibility**; • **Applied computing → Computer-assisted instruction**

Additional Key Words and Phrases: American sign language, education, feedback, user study

ACM Reference Format:
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 (April 2017), 30 pages.
DOI: http://dx.doi.org/10.1145/3046788

This material is based upon work supported by the National Science Foundation under award number 1400802, 1400810, 1462280.
Authors' addresses: M. Huenerfauth, B. Penly, S. Pillutla, M. Willard, and D. Hariharan, Rochester Institute of Technology, 152 Lomb Memorial Drive, Rochester, NY 14623, USA; emails: {matt.huenerfauth, bmp9681, sp6252, mtw2935, dh1723}@rit.edu; E. Gale, Hunter College, 695 Park Ave, New York, NY 10065, USA; email: egale@hunter.cuny.edu.

DOI: http://dx.doi.org/10.1145/3046788

1. INTRODUCTION

Based on survey and census data, researchers have estimated that over 28 million people in the U.S. are deaf or hard-of-hearing [Mitchell et al. 2006; Lin et al. 2011]. Several research studies have measured lower levels of written language literacy in deaf adults, which may be due to various language-exposure or educational factors. For example, in standardized educational testing studies conducted in the U.S., English literacy rates were found to be lower for deaf adults finishing secondary school, as compared to their hearing peers: the median literacy rate of deaf high school graduates was approximately at the fourth grade level [Traxler 2000] (fourth-grade corresponds to U.S. school students who are typically age 9 or 10).

Linguistics and educational research has established that the first few years of a child's life are a critical time for them to acquire their first language; children require adequate exposure to language, used in context, during this time [Mayberry and Eichen 1991]. When a child has (even mild) hearing loss, studies have found that this creates significant challenges for young children who are developing learning skills [Marschark et al. 2009].

One method of addressing the literacy issue mentioned above is to provide deaf children with exposure to language as early as possible. For instance, early access to American Sign Language (ASL) has been found to benefit learning of a second language (English) [Strong and Prinz 1997; Wilbur 2000], and some research studies have indicated that deaf children of Deaf[1] parents tend to read better than deaf children of hearing parents [Spencer and Lederberg 1997]. However, this is a minority of deaf children: More than 80% of children who are deaf or hard-of-hearing are born to hearing parents [Gallaudet Research Institute 2011; Mitchell and Karchmer 2004], and many children have difficulties in communication when parents are not fluent in ASL [Goldin-Meadow and Mayberry 2001].

This article investigates technologies that may be useful for parents or caregivers of deaf children to help support their learning of ASL. There are benefits when parents and caregivers rapidly learn ASL in order to support language acquisition of a deaf child: Even if parents do not become perfect signers, deaf children can learn ASL to a level of fluency consistent with native signers [Singleton and Newport 2004]. Of course, new parents of deaf children are busy, and they may benefit from educational technologies to support their learning of ASL, especially if those technologies can provide them with additional time flexibility in practicing their skills [Weaver and Starner 2011].

In a joint research project between Rochester Institute of Technology (RIT) and City University of New York (CUNY), we are investigating the design of software to automatically analyze a video of an ASL student's signing and provide immediate feedback. Such a tool would allow students to practice ASL independently, whenever it is convenient for them. Students learning ASL often produce videos of themselves: ASL students submit videos of themselves as homework assignments, and instructors watch them and provide feedback. Of course, such feedback is not immediate. For this reason, we are investigating how to produce an automated tool to give *limited* feedback about aspects of a student's movements that are *likely to be* errors (see Section 5.2). To be clear: our software tool would not replace feedback from ASL instructors, and it would only be able to identify a limited set of errors that may occur in a student's video. Thus, we envision that this tool could benefit students by providing *immediate* feedback to them about their signing (with the tradeoff of being that this feedback is less sophisticated, and perhaps less accurate). Our key motivation is to produce a tool that would benefit busy students, such as new parents, but many other students may also

[1]We follow the widely held convention of using the capitalized term "Deaf" to refer to people who identify as members of the Deaf Community or Deaf Culture, and we use "deaf" as a more general term.

benefit: In fact, ASL has become the third most studied language at U.S. universities; the enrollments in ASL courses rose by 19% from 2008 to 2013 [Goldberg et al. 2015].

One aspect of this research is the development of the video-analysis algorithms necessary to analyze the videos of students; that line of research is currently ongoing. However, the focus of the research presented in this article is to determine ASL students' subjective judgments about receiving feedback on their signing through video and to evaluate some initial concepts for the system design. This feedback from end-users will influence the selection of features for the educational system, and as discussed in Section 7, the results of this study also have broader implications for how human ASL instructors might best provide feedback to students in their courses. This article presents a pair of "Wizard of Oz" style studies to explore two questions:

Q1. Do ASL students report a subjective preference for using a tool that provides them with feedback about their signing, compared to just re-watching a video of their own signing?
Q2. Are students able to understand the feedback presented by such a tool, as measured by whether they are able to produce a new ASL video in which they correct the errors?

The first study (presented in Section 5) will focus on a comparison between three initial designs for how the system would present feedback for users, inspired by current technologies and practices of ASL students and instructors. A second follow-up study (presented in Section 6) will investigate variations of the most promising of the three designs from the initial study, to help further clarify how the feedback should be structured to be most useful for ASL students.

1.1. A Continuing Line of Research

This article is an extended version of a paper originally presented at the 2015 ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'15) [Huenerfauth et al. 2015]. That conference paper presented only our initial study (referred to as "Phase 1" in this article and described in Section 5), consisting of an evaluation of a Wizard-of-Oz prototype system displaying videos to ASL students of their own signing performance, with on-screen feedback provided in some cases. That initial study enabled us to evaluate three alternative designs for how to present feedback to students to communicate information about the errors in their signing.

While the Phase 1 study had identified a preference among students for videos of their signing that contained "pop-up" messages about errors they performed, that initial study did not investigate how such messages should appear. In feedback comments, participants in that study suggested that the on-screen feedback in our prototype could be improved in two ways: by simplifying the linguistic terminology used during the pop-up messages by using on-screen photos of "correct" ASL performances (to guide the ASL students how they should correctly perform aspects of ASL grammar). Thus, we subsequently conducted a follow-up study (referred to as the "Phase 2" study in this article) to evaluate and compare these additional design variations; details appear in Section 6. The results from this new study provide additional guidance about the appearance and presentation of corrective feedback messages to ASL students, which would not only shape the direction of our automatic feedback system, but may also be useful for ASL instructors in crafting effective feedback for students taking their courses.

In addition, this article provides more detail about our experimental study than was possible to include in the original conference paper: The Appendix to this article provides a comprehensive listing of the types of ASL errors that students made in their videos, which were the basis for the on-screen feedback shown (including the specific error messages shown to students for each type of error). Such information is necessary

for replication of our study. The Online Appendix for this article in the ACM Digital Library includes samples of the video stimuli presented in our studies.

1.2. Structure of this Article

The remainder of this article is structured as follows: Section 2 surveys prior work in language feedback tools for students learning ASL and explains the current methods that students in ASL courses received feedback on their signing. Section 3 outlines the specific hypotheses examined during the Phase 1 study, and Sections 4 describes the prompts and stimuli used in that study. Section 5 explains the methodology and results of the Phase 1 study. Section 6 describes our new Phase 2 study, which investigates additional design variations, and Section 7 presents our conclusions and future research directions.

2. RELATED WORK

There have been a variety of computational linguistic and speech technologies that have been used by researchers to create educational tools for students learning spoken/written languages; such technologies are surveyed in Hamidi and Baljko [2013]. For example, many of these systems help students to reduce their accent when speaking a language, or they may identify basic errors in their speech performance. A common thread among these systems is that they do not attempt to fully understand the entire message that is spoken by the student (since that would be beyond the state-of-the-art of their component technologies). Instead, these educational systems focus on a narrower task: They attempt to identify sounds or speech patters that are non-fluent. This design choice limits the types of errors that the systems identify; e.g., they don't generally identify semantic errors from an incorrect word choice. However, the systems are still useful to students: For example, they can inform them if an unusual consonant sound was spoken by the student, which may help the student notice an error in their accent.

In a similar manner, our research goal is to produce an educational tool for students who are learning ASL that can provide limited types of feedback to students about non-fluent elements of their movements, without the system understanding the specific semantic content of the ASL sentences being performed. Of course, the limitation of this design is that our software would not be able to identify all types of errors in student's ASL signing. For instance, the software would not know that a student made an error if she signed "SKY BROWN" (The sky is brown) instead of signing "SKY BLUE" (The sky is blue).

The advantage of designing an educational system that can operate without trying to fully understand the sentence produced by the student is that the software could be designed to function on input sentences that are unscripted. That is, the student would not be limited to performing sentences with a restricted vocabulary or sentences that follow a pre-determined script, which would be necessary to a achieve high accuracy in automatic ASL recognition software using state-of-the-art techniques, e.g., see survey in Cooper et al. [2011].

While a goal of our on-going research project at RIT and CUNY is to develop software that could automatically analyze the signing movements of a student from an analysis of the student's video, this article focuses on two studies that examine the user-interface and usability issues of the system design. Since this article does not focus specifically on the video analysis aspect of our research, prior research on sign-language recognition technology is not a focus of this related work section.

2.1. Automatic Analysis Education Tools

This section describes prior work that has utilized automatic video analysis technologies to recognize aspects of a student's movements, to create educational software to benefit students learning sign language. Some research has specifically focused on the

task of dictionary lookup for sign language. While languages with written forms use alphabetic or character-stroke-based methods for sorting words in a dictionary, the process of searching for an unfamiliar word in a dictionary resource is more complex for a language without a written form. In particular, students in ASL courses often have difficulty "looking up" signs in an ASL-English dictionary to find translations of particular words. Most ASL dictionaries are indexed according to the English translation of the ASL sign, but if the student does not know the meaning of the signing, finding the word can be difficult. Wang et al. [2010] implemented a system to allow users to search for words in an ASL dictionary by performing an ASL sign into a video camera; their recognition system returns a set of possible matches in the dictionary for the sign performed.

Other researchers have focused on methods to allow ASL students or instructors to find more video content of sign language, which could be used for classroom instruction or recognition practice by students. For instance, Monteiro et al. [2012] have developed a system for automatically searching video-sharing websites to identify videos that contain sign language.

Recognition technology has also been incorporated into educational games: For instance, the CopyCat system [Henderson et al. 2005] was a game in which the ASL student performed ASL signs or phrases (from a small predefined list that the system could recognize) while wearing special colored gloves with accelerometers. Players of the game attempted to give commands (using ASL signs or short ASL phrases) to an onscreen cat character. When the signs or phrases were correctly identified, thereby indicating to the student that they had performed them accurately, the system responded by causing the onscreen cat to obey the command. To enable the creation of a functional system despite limitations in the accuracy of sign language recognition technologies, the researchers selected a small set of signs or phrase-templates that the system could recognize. In contrast, the goal of our research is to create an educational tool that would allow students to receive feedback on any ASL signs or sentences that they perform. For instance, ASL instructors at different universities may assign homework assignments on different topics to students, and ideally the students could "rehearse" their ASL sentences by performing them for our system, to get feedback. Thus, the user-interface of the CopyCat system (controlling an onscreen character using ASL commands) was not well-suited to our research goals.

In subsequent research at Georgia Tech, [Weaver and Starner 2011] conducted interviews with parents of deaf children to understand their needs in ASL-learning technology. Based on that work, the researchers created software for mobile phones to display videos of signs, and students engaged in a "quiz" game in which they had to answer questions to identify the word in the video [Weaver et al. 2010]. However, that software did not provide feedback to users on the accuracy of their own ASL signing; the game required the student to correctly identify the sign shown on the screen of the device.

2.2. Education Tools without Automatic Analysis

Beyond the computerized educational tools discussed above, we also examined the current practices and technologies used by ASL students and educators. Our goal was to inform our design of software that could provide automatic feedback to students about their signing. The list below is arranged according to the amount of scheduling flexibility that the student would have in practicing and receiving feedback. Several of these items served as inspiration for experimental conditions examined in the Phase 1 study, described in Section 3.1.

—*Method #1: Receiving live feedback directly from an ASL instructor during a classroom interaction or private one-on-one meeting.* While receiving feedback live from a human expert would provide the student with the highest quality of feedback, this

method requires a substantial amount of ASL instructor time per student and must be scheduled in advance. One option is for students to schedule a private tutor for one-on-one live instruction and feedback; of course, this option may be prohibitively expensive for most students. As a more affordable option, students may enroll in classroom-based instruction with multiple students present. In that context, the ASL instructor must balance their time, providing feedback to all of the students in the room. The instructor focuses their attention on one student at a time during the classroom setting, as they circulate through the classroom during some type of in-class activity, e.g., the instructor may ask student to "pair up" and sign to each other while the instructor observes.

—*Method #2: Receiving feedback from an ASL instructor in a classroom setting using a recording device in the classroom.* One limitation of traditional classroom-based ASL instruction is that during in-class activities, students may perform errors when signing, but if the instructor is not looking at the student at that specific moment, there is no feedback provided to the student. Elaine Gale (co-author) teaches ASL at the university level, and students in her ASL classes have made use of mobile devices equipped with a recording application. Students can make recordings of each other during in-class activities. The key benefit of the use of these mobile devices is that the students can re-record their performance multiple times (if they notice mistakes in their own video) and then show the video to the instructor when they are ready to receive more sophisticated feedback. Of course, the instructor must still circulate through the classroom to watch the videos on students' devices, which can be challenging with often over 15 students in a classroom. From a time-flexibility perspective, students must still schedule themselves to attend class, but the instructor can provide feedback to the students in the classroom asynchronously, which can be more efficient.

—*Method #3: Students record themselves outside of a classroom setting and simply their own video.* This method is similar to the above technique, but the student makes a self-video performing sign language (outside of a classroom setting) and watches the video to look for errors. This alternative is essentially a time-shifted version of a classic "signing into a mirror" method for practicing, which ASL students traditionally employ to watch their own signing. While there is certainly greater time-flexibility with this self-video option, the downside is that the quality of feedback is much lower (since it is up to the student to look for errors in their own signing).

—*Method #4: Students record themselves outside of a classroom setting and submit their video to the instructor who provides a written evaluation or grade.* This method is the most common method that ASL instructors employ for providing formal feedback to students in their courses. It is common for "homework assignments" in an ASL course to consist of the following: The ASL instructor assigns a topic or prompt for the students, and students must record themselves performing ASL (in response to this assignment or prompt). The students submit the video of themselves to the instructor for grading. Later, the instructor provides a numerical grade and written comments. From a time-flexibility perspective, the student has flexibility about when to record their video, but the feedback received is not immediate. While the instructor is able to provide sophisticated comments, one challenge for the student is determining how the written comments from the instructor relate to specific moments in time in their videos.

—*Method #5: Students record themselves outside of a classroom and submit their video to the instructor, who uses a video-annotation tool to provide feedback that is time-synchronized to the student's video.* The prevalence of mobile devices capable of recording video has led to a new wave of software focusing on video recording and analyzing human performances. Specifically, for sports coaches and athletes who are

perfecting a complex physical movement, it has become increasingly common to use of video analysis software, e.g., Coach's Eye,[2] in which an instructor or coach watches the video of the athlete. Such tools enable the instructor to replay the video and add annotations (e.g., arrows, lines, text boxes) that appear a specific moments of the video. A similar tool is sometimes used for sign language instruction. For example, in several ASL courses offered at RIT, a commercially available web-based service called GoReact[3] has been used to enable students to submit videos of themselves signing and allow the ASL instructor to watch and provide time-synchronized annotations to the student as feedback. When the student watches their own video later, these feedback messages "pop up" during the video, with the feedback comments from the instructor. While time-based feedback may be easier for students to "match" to elements of their performance, the downside is that producing time-based annotation of ASL video is a time-consuming activity, even for experts [Lu and Huenerfauth 2014].

3. TWO STUDIES: PHASE 1 AND PHASE 2

Given the context of a student enrolled in an ASL course who would like to have additional feedback on their signing or practice their signing with a more flexible schedule, the goal of our research is to produce a tool that would allow such a student to perform some ASL in front of a video camera and receive automatic and immediate feedback. One use-case might be students using such a tool as a "spell check" (metaphorically speaking) while they are preparing to submit an ASL homework assignment video. Students may "rehearse" their ASL video, with the assistance of this educational tool, prior to performing a final version of their video, which the student would submit to the ASL instructor. (And the instructor could provide sophisticated and accurate feedback to the student.) In this article, we specifically focus on how to best present on-screen feedback to students on their ASL signing.

The research described in this article is organized into two research studies conducted on the RIT campus, with students enrolled in ASL courses as participants, to investigate and compare alternatives for how a feedback video could be shown to a student. In these studies, we investigated how to convey information to students about what was correct or incorrect about their ASL signing, and we have measured students' subjective preferences for different types of feedback (and how their signing improved after viewing feedback of different types).

Figure 1 contains a diagram that summarizes the two major studies ("Phase 1" and "Phase 2") in this article, along with presenting some additional terminology used to refer to the various components of each study: Specifically, the Phase 1 study included two rounds of video-recording sessions with students, referred to as "Recording Session A" and "Recording Session B." The Phase 2 student included one round of video recording with students, referred to as "Recording Session C."

3.1. Phase 1: Conditions and Hypotheses

Based on the final three methods of students receiving feedback listed in Section 2.2, we have identified three design alternatives below, which we evaluated in Phase 1:

1. **VIDEO:** We evaluated a lower-baseline condition, in which we simply replay the student's video for them to watch. This corresponds to Method #3 in Section 2.2, in which a student simply re-watches their own video.

[2]https://www.techsmith.com/coachs-eye.html.
[3]https://goreact.com.

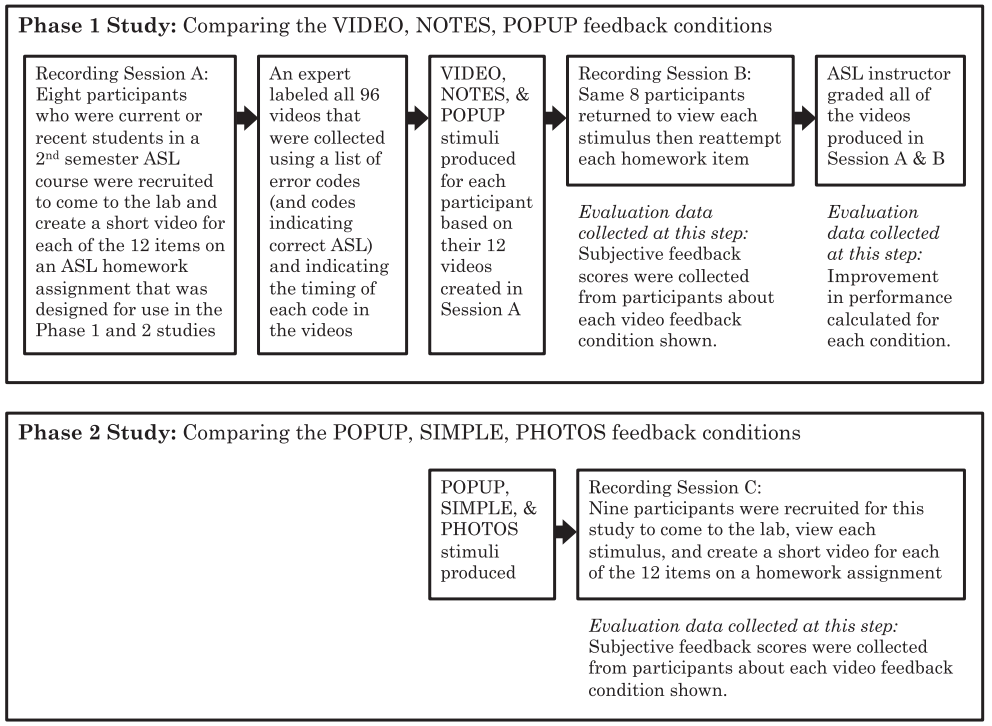


Fig. 1. Graphical illustration of the two studies presented in this article.

2. FEEDBACK: For comparison, we present videos that provide feedback to the student on their signing, under two *sub-conditions*:

2.a. NOTES: After analyzing a student's video, the system displays a written message to indicate any errors it identified in the video (or correctly performed linguistic elements that the system noticed). This corresponds to Method #4 in Section 2.2, in which an ASL instructor provides written feedback to a student about the video they submitted for grading.

2.b. POPUP: After analyzing a video, the system replays it to the student with messages that appear ("pop up") on the video when linguistic phenomena or errors are detected. At the end of this video, the student is provided written messages like those in the NOTES case, to summarize the errors. This corresponds to Method #5 in Section 2.2, in which an ASL instructor analyzes a student's video and provides time-synchronized feedback messages.

The NOTES sub-condition is similar to the VIDEO condition, but it is followed by text feedback about possible student errors. The POPUP sub-condition is similar to the NOTES sub-condition, with the addition of pop-up feedback messages during the student's video, about specific errors that may be occurring at that moment of the video. Section 5.3 describes how videos for each of these feedback conditions were produced to serve as stimuli during the Phase 1 study.

In the context of a student viewing a feedback video in one of these three conditions, prior to recording themselves performing ASL (e.g., as part of a homework assignment), we hypothesized that:

Hypothesis H1a: Students will report a subjective preference for videos that provide feedback about their signing (FEEDBACK), as compared to just a video of themselves signing (VIDEO).

Hypothesis H1b: Students will report a subjective preference for videos that provide time-synchronized feedback about their signing (POPUP), as compared to non-synchronized (NOTES).

Hypothesis H2a: When an instructor evaluates students' ASL performances "before" and "after" they viewed a video, there will be a significantly greater improvement when the video provides feedback about their signing (FEEDBACK), as compared to just a video of themselves signing (VIDEO).

Hypothesis H2b: . . . significantly greater improvement when the video provides time-synchronized feedback about their signing (POPUP), as compared to non-synchronized (NOTES).

Note that each of the "a" hypotheses compares the FEEDBACK condition to the VIDEO condition, and the "b" hypotheses compare the POPUP condition to the NOTES condition. Hypotheses "1" focus on subjective preferences reported on scalar response instruments; hypotheses "2" focus on the scores that an ASL instructor assigns to recordings of the students' signing.

3.2. Phase 2: A Brief Preview of Our Second Study

This section provides a brief preview of the Phase 2 study, to enable the reader to have a "big picture" view of the research presented in this article. Additional details about this Phase 2 study, including the specific hypotheses examined in that study and the results, will be described in Section 6.

At the conclusion of the Phase 1 study, as described in Section 5.6, we determined that the "POPUP" condition was the most preferred form of feedback among participants. In our subsequent Phase 2 study, we investigated alternative sub-variations of the POPUP condition to determine more precisely how the feedback on the screen should appear. These variations were inspired by participant feedback comments gathered during the Phase 1 study. Specifically, the following three conditions were examined in the Phase 2 study:

POPUP: This feedback condition was identical to the POPUP condition presented in the Phase 1 study.

SIMPLE: This feedback condition was identical to the POPUP condition, except that some of the linguistic terminology used in the on-screen messages was further simplified.

PHOTO: This feedback condition was identical to the SIMPLE condition, except that photos were added to the on-screen feedback to illustrate specific aspects of ASL grammar.

4. DESIGN OF COLLECTION PROMPT USED IN PHASE 1 AND PHASE 2 STUDIES

For the two studies presented in this article, we wanted to simulate the experience of an ASL student who must produce a video of themselves signing, as part of a homework assignment submission in an ASL course. Thus, we needed to provide our participants (who were ASL students) with a prompt similar to a "homework assignment" they might encounter in a second-semester ASL course.

Using the terminology often employed by ASL instructors, our assignment was an "expressive" assignment, which means that it was an assignment in which the student is asked to produce a video of themselves signing ASL in response to some topic or to provide an ASL translation of some English sentences. (Some ASL homework

assignments are "receptive," which require the student to view a video of ASL and attempt to understand the content.)

4.1. ASL Linguistics Background and Common Student Errors

ASL is a natural language conveyed through movements of the hands, body, head, eyes, and face [Valli et al. 2011]. Linguistics researchers studying second-language acquisition have observed that adults learning ASL often make errors involving movement, handshape (the way in which the fingers are bent to produce a specific hand configuration), and facial expression [Rosen 2004]. Therefore, for this homework assignment prompt, we wanted to encourage a wide variety of ASL linguistic features to be performed by students in a short video segment (to enable us to efficiently collect a wide variety of linguistic errors from the students).

As a brief background for the reader, this Section briefly summarizes some linguistic aspects of ASL. This information is relevant to the design of our homework assignment prompt used in our Phase 1 and Phase 2 studies (see Section 4.2), and much of this terminology is used in Section 5.2, including Table I and Table IV.

In ASL movements of the eyebrows, eyelids, nose, mouth, and head tilt are essential to conveying the meaning of sentences and other key grammatical information; these movements are sometimes referred to as "non-manual expressions" or more colloquially as "facial expressions" (although they include movements of the entire head). As an example, the ASL sentence "FATHER ARRIVE" (English translation: "Father arrives") can mean "Did father arrive?" with "YN-Question" facial expression performed during the sentence. Specifically, if the signer's head tilts forward with the eyebrows raised, this indicates that the sentence should be interpreted as a polar ("yes or no") question.

As discussed in Neidle et al. [2000], in addition to YN-Question mentioned above, other facial expressions used during ASL include, e.g.:

—*Negation*. The signer shakes their head from side-to-side to indicate that a sentence has an opposite meaning. The use of the ASL sign "NOT" is optional, but when the word "NOT" (or other negative words) are used during a sentence, the Negation facial expression is required.
—*WH-Question*. The head tilts forward with the eyebrows furrowed to indicate an interrogative who/what/where/why/when/how question.
—*Topic*. The eyebrows raise and head tilts back to indicate a topic introduced at the beginning of a sentence.
—*Conditional*. The head tilts back and to the side with the eyebrows raised to indicate the use of a time-phrase or an if-condition at the beginning of a sentence.
—*Rhetorical.* The head tilts back and to the side with the eyebrows raised during a phrase containing an interrogative word (e.g., "WHO") to indicate that the question should be interpreted rhetorically.

A common error among ASL students is to accidentally omit the facial expression during a sentence; in some cases, this omission can be easily detected: For instance, if the sentence includes an interrogative-question word ("WHO") then there should be a WH-Question facial expression used during the sentence.

During ASL, a signer will point to a location in three-dimensional space around their body to associate someone or something under discussion with that location [Valli et al. 2011]. To refer to that entity again, the signer (or their conversational partner) would simply point to this location, instead of saying the name of the entity again. (Many characterize this mechanism of ASL as a form of "pronoun" usage.) When pointing to one of these locations in space, typically the signer will use their eye-gaze to look at the locations, especially when they are establishing a new pointing location [Neidle et al.

2000]. ASL students may accidentally fail to use their eye-gaze appropriately when pointing in space in this manner.

When conveying proper names, titles, specialized terminology, or other specific categories of concepts, an ASL signer may use fingerspelling [Valli et al. 2011]. This is a method by which the signer performs a rapid sequence of handshapes, with each handshape corresponding to an individual letter of the alphabet. There are specific rules for proper hand location, hand orientation, and movement when someone is performing fingerspelling, and an ASL students may fail to obey these requirements.

4.2. Design of the Homework Assignment Prompt

As part of a prior research project, during which our laboratory collected video recordings of ASL, we develop techniques for designing "prompts" that could encourage someone being video-recorded to perform sentences in ASL with specific linguistic properties of interest. As part of that work, we measured the types of linguistic constructions that ASL signers tended to perform, in response to various types of prompt [Lu and Huenerfauth 2014]. This prior experience was valuable in designing a "homework assignment" prompt given to the students who participated in our Phase 1 and Phase 2 studies, with a goal of eliciting a wide variety of ASL linguistic constructions (that consequently would lead to the students making a wide assortment of errors in the video recordings). In addition, one of our co-authors is an experienced ASL instructor, who has been teaching ASL for nearly two decades.

Based on prior homework assignments used in her courses (and based upon an examination of sample assignments from other university ASL programs), we designed a homework assignment at an appropriate level of difficulty for a student who had recently completed approximately 1.5 semesters of university-level courses in ASL. We also examined frequently used textbooks and curricula: For instance, *Signing Naturally*[4] is a popular ASL curriculum used in U.S. universities, including at RIT, where we would be recruiting participants for this study. We therefore consulted the instructor's resource guide for the first two semesters in this curriculum series when designing our homework prompt, to ensure that it would match the skills of students. The homework was structured as a translation task, in which students were presented with twelve English prompts, and the student was asked to translate each into ASL and perform it in a video:

1. My house is green; your house is blue. My house is big; your house is small.
2. Her animal isn't a brown cat. What is her pet?
3. Is your favorite animal dogs or cats (which)?
4. In high school, your favorite class was not Spanish. Now, you're an ASL student.
5. You're not her cousin. Do you know her family?
6. Chocolate, I hate. Cookies, I love. What's your favorite food?
7. Her mother is reading a book. Her father is sitting. What are you doing?
8. His children are not here. Where does your family live?
9. Now, his father likes computers. He has no books.
10. He's my sister's son. Who is your son?
11. She grew up in France. Now, her parents live in England.
12. I saw her mother in the library. She had no glasses. How does she read books?

Prior to using this homework assignment prompt in any of our studies, we conducted two pilot testing sessions with ASL students who had 1.5 semesters of experience. These pilot sessions allowed us to determine whether our users would be able to understand this assignment, whether the instructions and terminology were clear, and whether

[4]http://www.dawnsign.com/series/signing-naturally-series/5.

students could complete this assignment in a 70-minute experiment session (including the time needed for them to prepare and record their responses). The results from the pilot tests indicated that the assignment and our recording protocol (below) were sufficiently well-specified to begin the study.

5. PHASE 1 STUDY

In order to evaluate the four hypotheses introduced in Section 3.1, we conducted a study that asked ASL students to video videos in the VIDEO, NOTES, and POPUP conditions. While our research goal is to build a software tool that could automatically analyze a video of ASL signing submitted by a student, the development of the video analysis components of the system are still ongoing. Therefore, in order to evaluate the different feedback conditions in this study, we used a Wizard-of-Oz style approach, as follows:

1. Our participants (ASL students) came to the laboratory, viewed the homework assignment prompt, and recorded a video of themselves signing each of the 12 items on the homework. We refer to this as "Recording Session A," and details appear in Section 5.1.
2. A human expert viewed the videos from the students and identified ASL linguistic errors in each video; the expert made notes about the category of each error and when it occurred during the video. A list of error types and details about this process appear in Section 5.2.
3. Based on the expert's notes, we used video editing software to produce a set of 12 feedback videos for each student (one video for each of their 12 recordings for the homework items, with four videos in each of the three conditions VIDEO, NOTES, or POPUP). Images of each condition and other details appear in Section 5.3.
4. The same set of participants individually returned to the laboratory on another day to view each of the feedback videos, immediately prior to making a new video recording of themselves for each of the 12 items on the homework assignment. We referred to this as "Recording Session B," and it is described in Section 5.4. At the end of the session, students responded to a questionnaire about their subjective impressions of the feedback video.
5. An ASL instructor evaluated each student's "before" and "after" homework videos and assigned a score to each video, based on specific linguistic elements of their performance, as described in Section 5.5.

The Phase 1 study used a within-subjects design: Since the homework assignment consisted of the student making 12 short videos in response to 12 short prompts, each participant saw four prompts in each of the three conditions (VIDEO, NOTES, or POPUP) during Step 3 above. Assignment of conditions to prompts (and the order in which they were presented) was counterbalanced.

5.1. Collection of Student Videos (Phase 1, Recording Session A)

Members of our research team posted recruitment advertisements for the study through email messages sent to students at RIT who were currently enrolled in their second semester of ASL courses (or who had recently completed their second ASL course). Our eight participants consisted of six women and two men, and all of them were RIT students pursuing undergraduate degrees. The students ranged in age from 19 to 24, and five of the participants had taken their first ASL class within 1 year of participating in the study. Although two of the participants were hard-of-hearing, none of the participants had deaf family members, and none of the participants grew up in a household that used ASL.

After participants completed the demographic questionnaire and informed consent documents, they were given approximately 5 minutes to read through the homework assignment prompt (Section 4.2) and to ask clarifying questions about the instructions for the assignment. The researcher conducting the recording session instructed the participants to approach this assignment as if it had been given to them as a homework assignment for an ASL class that they were taking. After the participant indicated that he or she understood the assignment, the researcher allowed them to work along for approximately 20 minutes to prepare. (Students in ASL classes will generally plan what they are going to say, before they make a video recording for a homework assignment.) The students were given access to the National Technical Institute for the Deaf online video dictionary,[5] a student edition of the *Signing Naturally* textbook, and other hardcopy ASL dictionaries.

As soon as they were finished with this preparation, the participants began recording the sentences. Participants sat on a stool 1.5 meters from a video camera on a tripod, which was at a height of 1.5 meters. The camera was connected to a Windows computer running recording software that saved each video (at 1920×1080 resolution at 30 frames-per-second) into a designated folder on the computer.

During the recording process, participants signed each of the 12 individual items on the homework assignment, one at a time, with short breaks in between. The researcher started and stopped the recording software for each of the items. These short breaks allowed the participants to look at their notes (created during the preparation period) and to prepare themselves for recording the next utterance. The students were instructed to begin and end their signing with their hands on their laps so that it would be obvious when the signing began and ended. Participants were also reminded to look into the camera while signing and to keep their homework paper on a nearby stool out of the camera's view. (ASL instructors generally ask that students look into the camera when creating homework assignments and do not read from their notes while recording.)

The researcher waited until the participant indicated that they were ready to record an utterance, started the recording, and then told the participant to begin. Once the participant was done signing the utterance (putting their hands on their lap), the researcher stopped the recording and checked that the video file was saved properly. After each item, if participants were not satisfied with their recording, then they were given the opportunity to rerecord the utterance as many times as needed until they were satisfied. However, participants were *not* allowed to view their recordings (they had to decide if they were happy with their performance based on their own memory). The rationale for not allowing participants to see their recordings at this time was that viewing their own recording was later used as the VIDEO feedback condition in "Recording Session B," described in Section 5.4.

After the 12 recordings were completed, the participants were compensated and asked if they would be willing to return at a later date for "Recording Session B." Since all the "Session A" participants indicated that they would return for "Session B," minimal debriefing was done at this time (to avoid revealing to the participants the purpose of the study).

5.2. Analysis of Videos from Recording Session A

In the next step of the study, the students' videos from the Recording Session A were analyzed by a professional ASL interpreter who had over 100 hours of experience at analyzing videos of ASL to identify linguistic occurrences. The work of this expert was one of the Wizard-of-Oz phases of this study; specifically, the expert was simulating the

[5]https://www.rit.edu/ntid/dictionary/.

work of the future *automatic* video-analysis software (operating within the near state-of-the-art of ASL recognition, see Cooper et al. [2011]). To keep the expert's activities within realistic limits of the future automatic system, the expert was only permitted to identify specific categories of errors (on a short list of error types that we expected our future automatic system to be able to identify). In addition, the expert was asked to note the timespan in each video (beginning and ending frame of video) when the error occurred.

As we explained in Section 2, our educational tool would complement (but not replace) the high-quality feedback from a human ASL instructor, who could consider the meaning that the student intended to convey. Given these limitations, we expect that our system would be able to identify a specific set of linguistic errors, for example, including some of the following:

—A student has incorrectly omits (or performs at the wrong time) an ASL facial expression for conveying questions, topics, negation, or time-phrases, which are linguistically required, as discussed in Section 4.1.
—The student has a lack of smoothness of movement, as characterized by overly large movements, stilted or halting movements, etc.
—The student uses incorrect hand location or orientation during fingerspelling.

Notably, since we do not expect that our automatic system would fully understand all of the signs performed by the student, our list of error codes does not include any semantic or word-order errors. Table I contains a brief listing of some of the error codes used in this study, and a longer listing appears in Table IV in the Appendix.

Given the short time duration of the videos in this study, each video was labeled with exactly two codes. In 18% of the videos collected during Recording Session A, the expert was not able to identify two *errors* in the student's video that matched our set of codes. In this case, the expert labeled a correct aspect of the video, using a pre-determined set of "correct codes" (see codes with names beginning with "correct" in Table I or Table IV) so that every video was annotated with two codes total. We anticipate that our future system would also provide feedback about specific aspects of a video that were correctly performed.

Because no feedback information would be provided for recordings in the VIDEO condition, it was actually not necessary to identify specific linguistic errors in the recordings for the purpose of producing stimuli (Section 5.3). However, in support of later "scoring" of videos (as described in Section 5.5), it was necessary to identify two linguistic properties of each video that would be a focus of this later scoring process.

5.3. Preparation of Stimuli to be Shown during Recording Session B

In order to further simulate, in a Wizard-of-Oz style, the operation of an automatic system for providing feedback to students, a member of our research team used video editing software to produce individualized feedback videos for each of the 8 participants from Recording Session A. The input to this process was the set of videos from Session A (Section 5.1) and the error codes and timing (Section 5.2) provided by the expert. With this information, the researchers produced video stimuli in three conditions to be shown to participants during Recording Session B (Section 5.4).

As we mentioned previously, for each participant, their 12 short videos were partitioned into three groups (4 per group). Each group was assigned to different condition (VIDEO, NOTES, POPUP). The assignment was determined prior to participant recruitment, ensuring videos were balanced across conditions and that conditions were presented to participants in a rotated order.

To produce the stimuli for the VIDEO condition, we simply re-used the video recordings collected during Recording Session A, without any modification. Figure 2 shows a

Table I. Brief Listing of Examples of Error Codes and Correct-Usage Codes Used for Analyzing Students' Recordings

Code	Description	Onscreen message displayed to student in FEEDBACK conditions (NOTES and POPUP)	Where the arrow should aim for POPUP stimuli
error_topic_beginning	Did a Topic facial expression begin too far from a clause boundary?	TOPIC facial expressions must start at the beginning of the clause.	Face
error_point_gaze	Is the signer performing a pointing sign without eye-gaze aimed at that location?	The first time you point to a location in space (to refer to someone or something), your eyes should look at the location.	Face
error_neg_lexical	Is the signer performing a negative word without a Negation facial expression?	You performed a negative word, but there was no Negative facial expression.	Face
error_hands_down	Is the signer putting their hands down too frequently during the signing?	You should try to sign in a fluid manner, without putting your hands down too frequently.	Torso
error_fingerspell_loc	Is fingerspelling happening now with the hand in an inappropriate location?	During fingerspelling, your hand should be near your mouth (without blocking it) and at the height of your chin or shoulders.	Hand
correct_contrastive	Is contrastive role shifting (torso swivel) happening now correctly?	You appeared to use good contrastive structure (torso movement) during your signing.	Torso

A complete list appears in Table IV in the Appendix.

Fig. 2. Image from the stimuli shown during the VIDEO condition of the Phase 1 study; the stimuli consisted of the student's video from Recording Session A.

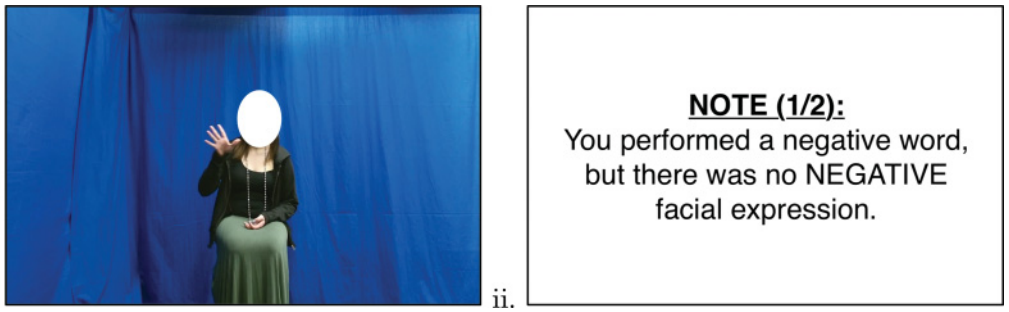


Fig. 3. Images from the NOTES condition in the Phase 1 study; the NOTES stimuli displayed the student's video from Recording Session A (i) followed by onscreen messages displayed after the video was finished.

screen image from one of these recordings. While the face of the participant has been obscured in the images shown in in Figures 2, 3, and 4, the face was clearly visible during the stimuli shown in our study.

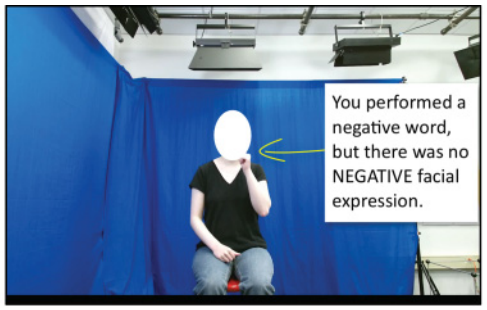
To produce the NOTES stimuli in our Phase 1 study, the researcher used video editing software to concatenate some still images containing text messages to the end of the videos from Recording Session A, illustrated in Figure 3. These notes at the end of the video included messages about the errors made during that particular sentence recording. We ensured that each text message was displayed for a sufficient amount of time (several seconds) and was in a large font size for comfortable reading. The text of this message was based on the error code (see Table I or Table IV), and it was not customized for the particular participant's video. In this way, the use of "fixed" error messages as feedback to the student was meant to simulate the type of feedback that we expect the future system to perform. In this way, the stimuli-creation work in this phase is another Wizard-of-Oz aspect of the study.

To produce the POPUP version of the stimuli, Camtasia Studio[6] was used to add pop-up messages to the videos from Recording Session A, to identity errors made during signing. Each message included an arrow pointing to part of the signer's body (e.g., face, hands, torso, etc.). The location on the body where the arrow pointed was based upon the error code, as show in Table I. The text of these pop-up messages was identical to the text used in the feedback presented at the end of the NOTES stimuli. The timing of the pop-up messages was based on the information provided by the expert (Section 5.2). As in the NOTES stimuli, to ensure that participants had sufficient time to read the messages, as well as observe the respective video segment, effects such as slow-motion and freeze-frame were employed. The font size was set large enough for comfortable viewing on the display monitor used during the Recording Session B (Section 5.4). In addition to the pop-up messages displayed during the student's recording, the videos in the POPUP condition also contained summarizing notes at the end of the video, identical to those shown in the NOTES condition. Figure 4 contains some images of the pop-ups shown during the video. Example stimuli from our study are available in the Online Appendix for this article on the ACM Digital Library.

5.4. Collection of Student Videos during Recording Session B

The same set of eight participants who had participated in Recording Session A were invited back to the laboratory to participate in Recording Session B. When they were at the laboratory during Recording Session A, many students had made hand-written

[6]https://www.techsmith.com/camtasia.html.

i. 

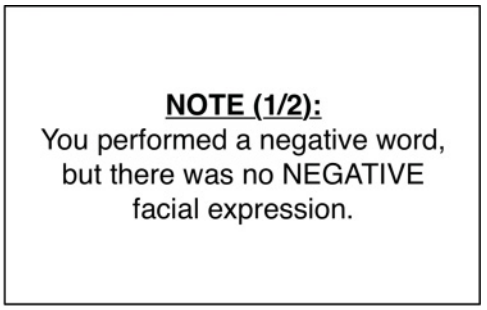
ii. 

Fig. 4. Images from the POPUP condition in the Phase 1 study; the POPUP stimuli displayed the student's video from Recording Session A with overlaid messages (i) followed by onscreen messages displayed after the video was finished (with content identical to the overlaid "popup" messages shown earlier in the video).

notes on the homework paper while they were planning their ASL sentences, and so these sheets were retained by the researcher and given back to the participant during this "Recording Session B." After completing another set of informed consent paperwork and briefly reviewing their homework assignment paper from Recording Session A, the participants were ready to begin recording.

The camera setup and distance was identical to that during Recording Session A, with one difference: a 27-inch LCD monitor was placed directly below the camera so that the participant could be shown videos immediately prior to their performing each ASL sentence to be recorded. Before re-recording each of the 12 short videos for the homework assignment, the researcher showed the participant a feedback video, in one of the three conditions: VIDEO, NOTES, or POPUP. To be clear, the videos shown to each participant showed that participant – that is, Participant "A" saw videos of herself from Recording Session A (with feedback added to the videos, depending on the condition), and Participant "B" saw videos of himself from Recording Session A. The video display monitor was turned on only to play the stimulus video and was shut off shortly after so that it would not distract the participants while they were making their recordings.

At the end of the recording session, participants answered 0-to-10 scalar response questions about their subjective impression of the quality of each type of stimuli (VIDEO, NOTES, and POPUP) presented during the study, with 10 representing the best quality. Participants were asked to assign a score to each of the three conditions in the study.

5.5. Evaluation of Videos from Recording Session A and Recording Session B

An experienced ASL instructor (second author) evaluated the student videos from Recording Session A and from Recording Session B. This instructor is a professor and director of a Deaf and Hard of Hearing Education program at CUNY Hunter College and has been teaching university ASL courses for almost two decades. She works in New York City and was not present for the recording sessions in Rochester (several hours away). Prior to sending her the videos, the research team at RIT randomly sorted and renamed the video files, to hide whether each video was collected from Recording Session A or from Recording Session B.

Rather than assigning an overall holistic score to each video, instead, our evaluation was *targeted* to focus on the specific linguistic properties that had been mentioned in the feedback messages in the stimuli. So, for each video, the instructor was asked to assign a score for two linguistic characteristics of the video. These

Table II. Subjective Evaluation Scores for Each Condition

		FEEDBACK	
Participant	VIDEO	NOTES	POPUP
"A"	10	10	10
"B"	3	7	9
"C"	6	7	7
"D"	1	5	8
"E"	5	7	9
"F"	3	9	10
"G"	3	5	10

properties were selected based on the error or correct-usage codes assigned to that video during the expert analysis (Section 5.2). Across all videos in the study, the ASL instructor was asked to assign a numerical score (on a 0-to-10 scale) for these linguistic aspects of some recordings (two aspects per video): Pointing, Hands-Up-When-Signing, Size-of-Signing, Topic, Wh-Question, Negation, Yes-No-Question, Time-Transition, Contrastive-Structure, Looking-at-Camera, Smoothness, and Fingerspelling. For example, if a video had been labeled with codes relating to the Topic facial expression (e.g., error_topic_beginning in Table I) and Pointing (e.g., error_point_gaze in Table I), then the instructor assigned a numerical score (on a 0 to 10 scale) for the Topic and Pointing linguistic properties for that video. To evaluate the Topic property, the instructor judged whether the signer performed the topic facial expression correctly, at an appropriate time. To evaluate Pointing, the instructor judged whether the signer pointed to locations in space when signing to represent entities under discussion, whether eye-gaze was used correctly during this pointing, whether the signer was consistent in the locations that were pointed to, and whether the signer used the correct hand shape for pointing, etc.

The instructor assigned scores of 9–10 for videos with all or almost all grammatical features correctly used, 7–8 when most of the features correctly used, 5–6 when some features correctly used, and 0–5 when few to no features correctly used (or missing). The average of the two grammatical features for the video was the final score. For example, if the instructor assigned a score of 8 for Topic and 6 for Pointing, then the video received a score of 7. To evaluate the degree to which a student's performance changed from Recording Session A to Recording Session B, we subtracted the score from the Recording Session A video from the score for the Recording Session B video. For instance, if the student's Session A video received an overall score of 7 and the Session B video (for that same homework item) received a score of 9, then the student had a 2-point improvement in the score.

5.6. Results and Comparisons from the Phase 1 Study

This section presents the results of our comparison of the three feedback conditions based on the participants' subjective preferences and the overall improvement in their signing during the Phase 1 study. Table II presents the responses from participants at the end of the Phase 1 study when they were asked to rate "the quality of this feedback, on a 0-to-10 scale (10 = best)" for each of the three conditions in the study: VIDEO, NOTES, and POPUP.

Hypothesis H1a considered whether students reported a subjective preference for videos that provide feedback about their signing (FEEDBACK), compared to a video of themselves signing (VIDEO). Median response for FEEDBACK videos (union of NOTES and POPUP videos) was 8.5, and the median for VIDEO was 3. The distributions in the two groups differed significantly (Mann-Whitney $U = 16.5$,

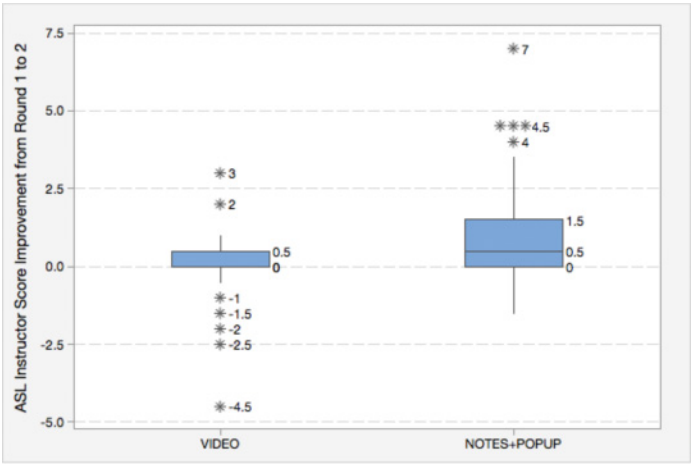


Fig. 5. Boxplots of improvements in ASL instructor scores for student's videos from Recording Session A to Recording Session B for the VIDEO condition and for the FEEDBACK condition (NOTES+POPUP).

$n_{FEEDBACK} = 16$, $n_{VIDEO} = 8$, $p < 0.05$). Thus, *Hypothesis H1a was supported:* Videos with feedback information received higher subjective scores from ASL students.

Hypothesis H1b considered whether students reported a subjective preference for videos that provide time-synchronized feedback (POPUP), compared to non-synchronized (NOTES). Median response for POPUP was 9.5, and for NOTES, 7. The distributions in the two groups differed significantly (Mann-Whitney $U = 11$, $n_{NOTES} = 8$, $n_{VIDEO} = 8$, $p < 0.05$). Thus, *Hypothesis H1b was also supported:* Students gave higher scores for POPUP.

Hypothesis H2a considered whether, when an instructor evaluates students' ASL performances "before" and "after" they viewed a video, there was a significantly greater improvement when the video provides feedback about their signing (FEEDBACK), as compared to just a video of themselves signing (VIDEO). Figure 5 displays these improvements in ASL instructor scores; the box indicates quartiles, whiskers indicate 1.5 inter-quartile ranges, and asterisks indicate outliers.

Hypothesis H2a was supported. A two-sample t-test, $t_{(df = 71)} = 3.49$, $p = 0.0008$. The scores for students' recordings showed a greater improvement when students were shown a feedback message, as compared to simply being shown the video of their ASL signing from round 1.

Hypothesis H2b was not supported. A two-sample t-test did not reveal a significant difference in the improvement of the scores for the student videos in the NOTES vs. POPUP conditions. Notably, students gave higher subjective scores for POPUP (H1b).

Informally, at the end of the Phase 1 study, some participants mentioned to the researcher that they didn't understand some of the linguistic terms used during the on-screen messages displayed in the NOTES or POPUP condition. We had originally selected the terminology used on those feedback messages (see Table IV) based on similar terms used in ASL curriculum commonly used in U.S. universities, but we can report anecdotally that some participants did not find the terminology understandable. In hand-written feedback comments collected from participants at the end of the Phase 1 study, three participants mentioned that the text of the feedback messages weren't specific or clear enough for them to understand. For example, one commented "Information was rather vague. Use less technical terminology." In addition, four participants requested a video or photograph to be shown of what they *should have* done

in order to correct the error; specifically, they requested some type of illustration of how they should perform the correct facial expression. For example, one participant commented "I didn't know what facial expression goes with a time/if sign. If an example or picture could be added to the feedback to show the proper facial expression that would be awesome."

6. PHASE 2 STUDY

As discussed above, some of the feedback comments from participants in the Phase 1 study suggested alternative design variations that would make the feedback of the POPUP stimuli more useful to students, such as:

—Alternative terminology for describing the ASL language errors to make it more understandable to students (e.g., avoiding linguistic jargon or giving specific direction as to how to improve the performance).
—Providing an image of a "correct" performance of specific types of ASL facial expressions, which are mentioned in corrective feedback comments, to make it easier for students to improve their performance.

To further investigate these design options, we conducted a follow-up study to display variations in the feedback messages along these lines. In this "Phase 2" study, we elicited subjective ratings (and feedback comments) from ASL students about the understandability of the messages and the students' subjective preference for three variations (which previously listed briefly in Section 3.2):

POPUP: This feedback condition was identical to the POPUP condition presented in the Phase 1 study.

SIMPLE: This feedback condition was identical to the POPUP condition, except that some of the linguistic terminology used in the on-screen messages was further simplified. We removed any complex linguistic "jargon," and we replaced it with simpler text.

PHOTO: This feedback condition was identical to the SIMPLE condition, except that photos were added to the on-screen feedback to illustrate specific aspects of ASL grammar. So, the complex "jargon" was removed, and, in addition, photographs were displayed onscreen showing correct performance of the facial expression, correct hand location during fingerspelling, or eye-gaze during pointing.

6.1. Hypotheses for Phase 2 Study

Our Phase 2 study allowed us to examine an additional pair of hypotheses related to the use of simplified language on feedback messages and the use of photo illustrations. Since during our Phase 1 study, Hypothesis H2b was not supported (we did not measure a difference in student performance in the POPUP vs. NOTES condition), we decided to focus on participant's subjective preferences in this Phase 2 study. To keep the nomenclature in this article consistent, we have numbered these additional hypotheses with "3."

Hypothesis H3a: Students will report a subjective preference for videos that use simplified terminology when providing time-synchronized feedback messages about their signing (SIMPLE), as compared to feedback using more technical terminology (POPUP).

Hypothesis H3b: Students will report a subjective preference for videos that include a photo illustration of correct ASL performance when providing time-synchronized feedback messages about their signing (PHOTO), as compared to messages with identical text but without photo illustrations (SIMPLE).

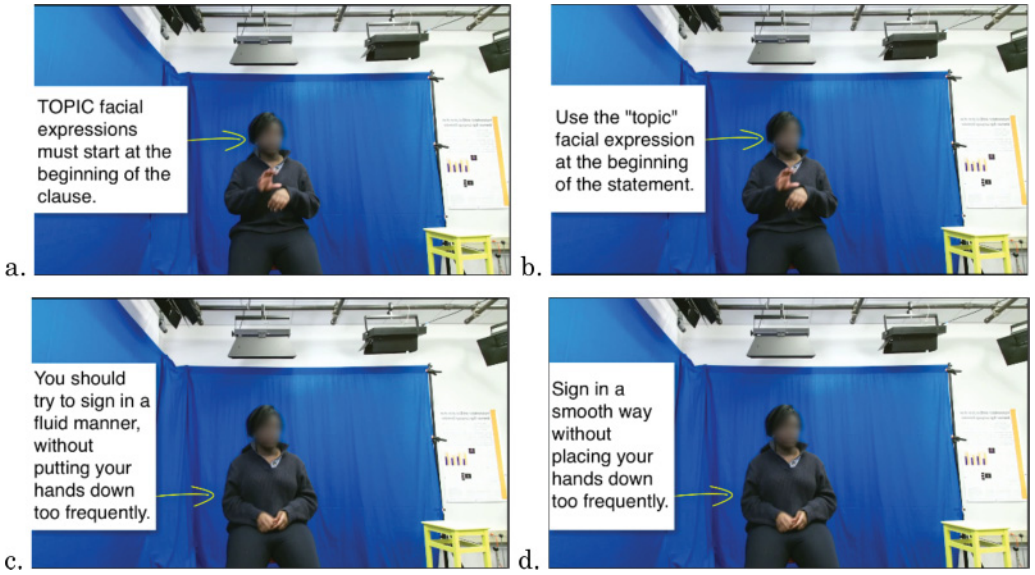


Fig. 6. Comparison of the text that appeared in onscreen messages from the videos in the POPUP (a, c) and SIMPLE (b, d) condition in the Phase 2 study, for error codes error_topic_beginning (a, b) and error_hands_down (c, d) as listed in Table IV.

6.2. Stimuli for Phase 2 Study

To produce stimuli for the Phase 2 study, we selected a set of 12 videos from a participant in the Phase 1 study, and we produced three versions (POPUP, SIMPLE, and PHOTO) of each of the 12 videos (for a total of 36 stimuli videos for use in the Phase 2 study).

For the POPUP stimuli, since four of these short videos had already been produced in the POPUP condition for use in the Phase 1 study, no additional work was required for those four videos. For the remaining eight videos, we added on-screen pop-up style messages to indicate the errors (or correct elements, if two errors could not be found) for each video, using the Camtasia Studio software, as was done in the Phase 1 study. Table IV in the Appendix contains details about the on-screen messages displayed and where the arrow graphic pointed for each pop-up message.

After having produced POPUP versions of all 12 short videos, we modified each stimulus to produce a SIMPLE version. The timing and graphical-style appearance of the onscreen messages was identical between the POPUP and SIMPLE conditions: The only difference was that the terminology of the text messages was somewhat simplified, with an effort to avoid any technical linguistic jargon. For example, an onscreen message mentioning a "clause" was adjusted so that it used the term "sentence" instead, which while not as precise linguistically, was sufficient for the onscreen message. Table IV in the Appendix contains a full set of text messages shown in the POPUP and the SIMPLE conditions of the study. Figure 6 displays some examples of the POPUP and SIMPLE conditions side-by-side for comparison; although the face of the participant is blurred in the photographs in this figure, the face was clearly shown during Recording Session C in the Phase 2 study.

After having produced the SIMPLE versions of all 12 short videos, we modified each stimulus to produce a PHOTO version. The timing and text content of the onscreen messages was identical between the SIMPLE and PHOTO conditions: The only difference was that a photograph was added to the screen that illustrated a correct ASL performance related to the topic of the onscreen message. For instance, if the onscreen

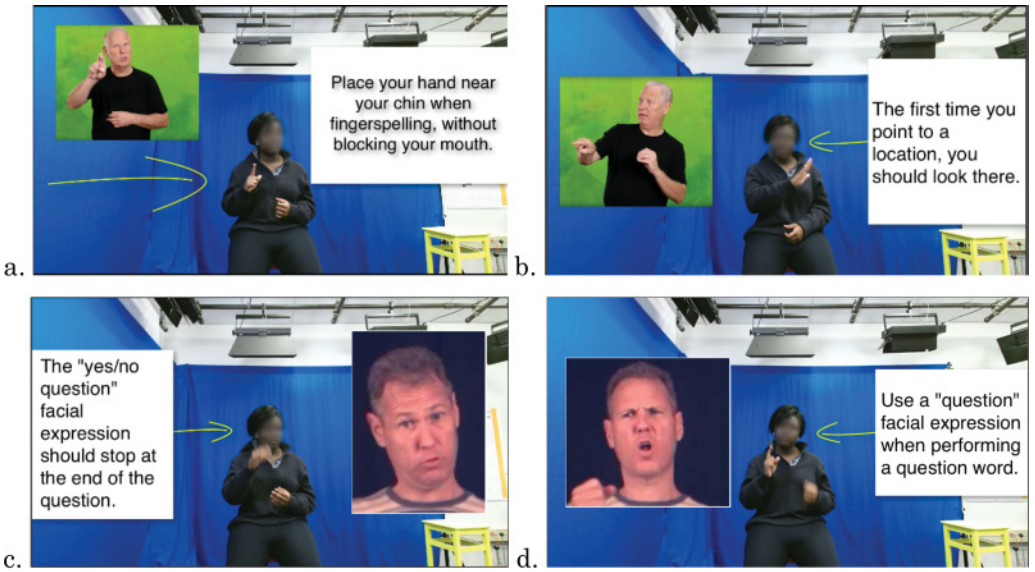


Fig. 7. Examples of onscreen feedback messages shown during the PHOTOS condition during the Phase 2 study, for error codes: error_fingerspell_location (a), error_point_gaze (b), error_ynq_end (c), error_whq_lexical (d) as listed in Table IV.

message mentioned a particular type of facial expression (e.g., Negation), then a photo of a fluent signer performing this facial expression was displayed on the screen. If the onscreen message mentioned using eye-gaze to look at the location where pointing or holding your hand in correct location/orientation when performing fingerspelling, then a photograph demonstrating proper eye-gaze or hand position was shown. Given the wide-screen proportion of the video image, there was sufficient space for this photograph to be overlaid on the video without blocking the human signer; typically the photo was placed on the opposite side of the human signer as the on-screen pop-up message, as shown in Figure 7. The demonstration images were drawn from photographs of fluent ASL signers appearing in guide materials from a linguistics research project at Boston University [MacLaughlin et al. 2000] and still images taken from an online video performance of a fluent ASL signer [Bahan 2016].

6.3. Recruiting Participants for the Phase 2 Study

In a similar manner to which participants were recruited for the Phase 1 study, our research team advertised for the study (conducted 1 year later) through email messages sent to students at RIT who were currently enrolled in their second semester of ASL courses (or who had recently completed their second ASL course). Our nine participants in the Phase 2 study consisted of six women and three men, and all of them were RIT students pursuing undergraduate degrees. The students ranged in age from 18 to 25. None of the participants had deaf family members. Seven of the participants described themselves as "hearing," and two, as "hard of hearing." These two hard-of-hearing students had only begun studying ASL within the past 2 years; they did not grow up using ASL.

6.4. Collecting Data during Phase 2 Study, Recording Session C

The methodology for the Phase 2 study was simpler than that used in Phase 1. Specifically, the participants only visited the lab on one occasion: They were provided with an

Table III. Subjective Evaluation Scores for Each Condition

Participant	POPUP	SIMPLE	PHOTOS
"Z"	8	8	9
"Y"	6	2	9
"X"	5	5	10
"W"	10	10	9
"V"	5	5	9
"U"	7	8	9
"T"	5	8	9
"S"	2	4	8
"R"	10	8	10

ASL homework assignment (same homework assignment as used in Phase 1), and they were given 20 minutes to prepare to record a video of themselves. The placement of the camera and the video display screen in front of the camera was identical to that used during Recording Session B in the Phase 1 study. Prior to the participant attempting to record each of the 12 items on the homework assignment, a video was displayed showing a student attempting that homework assignment with on-screen feedback messages displayed onto the screen in one of the three feedback conditions POPUP, SIMPLE, or PHOTOS. After viewing each stimulus, the participant was video-recorded while they attempted each of the 12 items on the homework assignment.

Similar to the Phase 1 study, each participant in Phase 2 saw videos in all three conditions: four videos in the POPUP condition, four in SIMPLE, and four in PHOTOS. Latin squares were used to assign conditions to each video that was displayed during the study, with each participant seeing videos 1–4 displayed in the same condition (POPUP, SIMPLE, or PHOTOS), videos 5–8 in the same condition, and videos 9–12 in the same condition. At the end of the study, participants were asked a subjective feedback question about each of the three feedback conditions.

Unlike the Phase 1 study, in which we subsequently had an expert ASL instructor score the students' homework submissions, our Phase 2 study focused only on participant's subjective preferences. We still asked the participants to attempt the homework assignment and produce video-recordings in order to simulate the experience of looking at these videos as part of the process of preparing and creating homework assignment videos; however, the videos produced by participants were not evaluated as part of this study. The rationale was that the results of Phase 1 did not reveal any difference in performance scores between NOTES and POPUP; so, we decided to focus only on subjective preference in the Phase 2 study.

6.5. Results of the Phase 2 Study

This section presents the results of our comparison of the three feedback conditions (POPUP, SIMPLE, and PHOTOS) based on the participants' subjective preferences during the Phase 2 study. Table III presents the responses from participants at the end of the study when they were asked to rate "the quality of this feedback, on a 0-to-10 scale (10 = best)" for each of the three conditions: POPUP, SIMPLE, and PHOTOS.

Hypothesis H3a considered whether students reported a subjective preference for videos that use simplified terminology when providing time-synchronized feedback messages about their signing (SIMPLE), as compared to feedback using more technical terminology (POPUP). Median response for POPUP videos was 6, and the median for SIMPLE was 8. No significant difference was observed in the distributions of scores (Mann-Whitney $U = 44.5$, $n_{POPUP} = 9$, $n_{SIMPLE} = 9$, $p > 0.05$). We were unable to reject the null hypothesis: *Hypothesis H3a was not supported.*

Hypothesis H3b considered whether students reported a subjective preference for videos that include a photo illustration of correct ASL performance when providing time-synchronized feedback messages about their signing (PHOTO), as compared to messages with identical text but without photo illustrations (SIMPLE). Median response for SIMPLE was 8, and for PHOTOS, 9. The distributions in the two groups differed significantly (Mann-Whitney $U = 14$, $n_{SIMPLE} = 9$, $n_{PHOTO} = 9$, $p < 0.05$). Thus, *Hypothesis H3b was supported*: Students gave higher scores for videos in the PHOTOS condition, as compared to the SIMPLE condition.

Some of the written feedback comments from participants collected at the end of the study helped to explain why the SIMPLE condition was not significantly preferred to the POPUP condition. Some participants felt that the wording used in the SIMPLE condition was less precise or less informative than that in the POPUP condition; for instance, Participant S commented, "The notes made were unclear as what the signer did wrong (not easy to understand wording)." Participant U said that the SIMPLE version of the videos "did not show how to fix the problem." Many participants reported difficulty in noticing the difference between the SIMPLE and POPUP conditions. Participant U wrote, "I could not tell the difference between the [POPUP] version B and [SIMPLE] version C." When comparing the POPUP and SIMPLE conditions, Participant V indicated that they "weren't very different from each other." Participant S said that they "felt the same and weren't very helpful."

Participants commented that the time-synchronized nature of the feedback messages was helpful, which is in line with the results of the Phase 1 study. For example, Participant V said, "I liked how it was times when mistakes happened." Participant U noticed that the speed of the video was often slowed when a feedback message was displayed, commenting, "I liked how it slowed down during mistakes and reviewed them at the end." However, some participants wanted to have more control over the video speed, e.g., Participant V commented "Make it easier to slow down feedback, it's hard to process ASL guidance as an inexperienced signer."

As indicated by the significant results presented above in regard to Hypothesis H3b, in general, participants liked the images in the PHOTO condition. In regard to the PHOTOS videos, Participant X said "I really liked the examples of how it should look done by a professional. Just knowing there's an error without seeing the correction is frustrating so that was great." In fact, some participants advocated for additional forms of visualization be provided to help illustrate the feedback, including video. Participant Y suggested that "video feedback showing how to do the sign or facial expression correctly would be even better than pictures," and Participant S recommended that we use "short video clips to explain the notes."

7. CONCLUSIONS AND FUTURE WORK

The goal of our research is to benefit ASL students by enabling them to review and evaluate their signing skills through automatic, immediate, outside-of-classroom feedback. Such a tool may provide students with additional time-flexibility (reducing dependence on arranging meetings), which may benefit busy students, such as new parents. Ultimately, the creation of new educational technologies for learning ASL may benefit deaf and hard-of-hearing individuals through the promotion of sign language education among parents, caregivers, or colleagues.

Our Phase 1 study compared three methods of presenting feedback, inspired by current approaches used by ASL students (Section 2), and we found that students gave higher subjective scores to videos with text-based feedback messages about their errors or correct-usage. As evaluated by an ASL instructor, the students who received feedback messages also showed greater improvements in their signing (when we compared videos from Recording Session A and B). This finding suggests that the set of

error and correct-usage codes used in this study were sufficient for providing beneficial feedback to the participants. Since these feedback "codes" were limited (they did not include errors based on the meaning of the student's video, they were not based on full-understanding of what they were signing, and they did not provide feedback messages that were specifically customized to the words of the student's video), this is a useful finding. It suggests that a future automatic video analysis system based on this type of feedback would be useful for ASL students.

In our Phase 1 study, participants gave higher subjective scores to those videos that presented the text-based feedback in a time-synchronized manner with their video. And in the Phase 2 study, we found that participants gave higher subjective scores to videos that illustrated the feedback with photographs of correct ASL performance. These findings suggest how our future educational tool should function: Specifically, it is useful for the system to identify the specific times during the video when errors occur, so as to support time-synchronized feedback for students. Further, it is important that the system show the student visually how to perform the ASL facial expression or movement correctly; for instance through images provided onscreen.

While the focus of our research project is the creation of an automatic educational tool for students learning ASL, in order to conduct the studies in this article, we had to simulate our future automated system. Thus, it was actually human "wizards" who produced the feedback for the student, in this Wizard-of-Oz prototype evaluation. Given that we were actually studying how students reacted to feedback that was generated by humans, our findings may also have implications for sign language educators or education researchers. Specifically, our findings may suggest how students would respond to various forms of communicating feedback information on their ASL performance, whether that feedback is produced by an automated system or by an ASL instructor. Given the increasing popularity of university courses on ASL [Goldberg et al. 2015], including online courses, understanding how to effectively convey feedback to ASL students electronically is increasingly important. Based on our results, it seems that students prefer time-synchronized feedback on videos of their signing, and they also benefit from visual illustration of how to perform aspects of ASL correctly when they correct these errors.

7.1. Limitations of this Research and Plans for Future Work

One limitation of the Phase 1 and Phase 2 studies was the relatively small number of users. While some statistically significant differences were observed (and the population size was above customary minimums for application of the statistical tests employed in the study), it would be valuable to reproduce this study with a large number of ASL students in future work.

An addition limitation in our Phase 1 study (which considered the improvement in participants' signing performance) was that a single ASL instructor performed the evaluation of the student videos. To address this, we made use of a blind-evaluation approach in which a highly experienced evaluator judged the video quality on a list of specific criteria, but, of course, in future work, we could replicate this study with additional ASL instructors independently evaluating the student videos.

A fundamental limitation of the two studies presented in this article is the risk that the Wizard-of-Oz enactment of our automatic feedback system may not accurately reflect the future system. To minimize this risk, we made use of a detailed table of error-codes, on-screen messages, and standardized appearance of feedback on the stimuli videos. However, the decision of where and when to display feedback was based upon an expert who viewed the video and determined when specific error codes applied to each video. In the future automated system, the software may make some errors when analyzing student's recordings; in a future study, it would be useful to examine whether

students' preferences for onscreen feedback differ when there are errors in the provided information. Further, in our Phase 1 study, we needed to ask students to return on a second occasion (Recording Session B) to provide them with a video showing feedback; in our future automatic system, the feedback would be more instantaneous than we were able to provide in this Wizard-of-Oz setting. Given these limitations, in future work, we plan to explore how students respond to feedback messages from an automatic system. Therefore, a major aspect of our future research will be our continuing efforts to design and develop software capable of automatically identifying errors in students' videos. Our research team at RIT is collecting and linguistically analyzing videos of ASL students and experienced ASL signers, is being provided to collaborators at CUNY who will be using this data to train computer vision algorithms for sign language recognition. When initial prototypes of this technology are ready, we intend to replicate the feedback studies presented in this article with the working system.

APPENDIX

Table IV. Complete list of the Error Codes and Correct-usage Codes Used For analyzing Students' Recordings during the Phase 1 and 2 Studies, with Information About the Onscreen Messages in the NOTES, POPUP, SIMPLE, and PHOTOS Conditions

Code	Description	Onscreen message text displayed to student in NOTES and POPUP conditions	Onscreen message text displayed to student in the SIMPLE or PHOTO conditions (Phase 2)	Where the arrow should aim for pop-up messages
error_cond_lexical	Did the signer perform a time or "if" word without a conditional facial expression?	If you want to mention a time or and if at the beginning of a clause, you should use a TIME/IF facial expression.	Use a "time/if" facial expression at the beginning of a sentence when mentioning a time or an "if" phrase.	Face
error_eye_camera	Did the signer fail to look at the camera appropriately?	Your eyes should look at the camera most of the time when making an ASL video.	Look at the camera most of the time.	Face
error_fingerspell_loc	Is fingerspelling happening now with the hand in an inappropriate location?	During fingerspelling, your hand should be near your mouth (without blocking it) and at the height of your chin or shoulders.	Place your hand near your chin when fingerspelling, without blocking your mouth.	Hand
error_fingerspell_handshape	Is fingerspelling happening now with the hand in an inappropriate handshape?	During fingerspelling, you should be careful to perform the hand shapes accurately.	Use ASL alphabet handshapes when fingerspelling.	Hand
error_fingerspell_movement	Is fingerspelling happening now with the hand in an inappropriate movement?	During fingerspelling, you should avoid unnecessary movements of the hand.	Your hand should remain near your chin when fingerspelling, without blocking your mouth.	Hand

(Continued)

Table IV. Continued

Code	Description	Onscreen message text displayed to student in NOTES and POPUP conditions	Onscreen message text displayed to student in the SIMPLE or PHOTO conditions (Phase 2)	Where the arrow should aim for pop-up messages
error_hands_down	Is the signer putting their hands down too frequently during the signing?	You should try to sign in a fluid manner, without putting your hands down too frequently.	Sign in a smooth way without placing your hands down too frequently.	Torso
error_hands_large	Is the signer using overly large or exaggerated movements?	It is not necessary to make large or exaggerated movements.	Your signs should not be too big.	Torso
error_neg_lexical	Is the signer performing a negative word without a Negation facial expression?	You performed a negative word, but there was no NEGATIVE facial expression.	Use a "negation" facial expression when performing a negative word.	Face
error_point_gaze	Is the signer performing a pointing sign without eye-gaze aimed at that location?	The first time you point to a location in space (to refer to someone or something), your eyes should look at the location.	The first time you point to a location, you should look there.	Face
error_topic_beginning	Did a Topic facial expression begin too far from a clause boundary?	TOPIC facial expressions must start at the beginning of the clause.	Use the "topic" facial expression at the beginning of the statement.	Face
error_whq_lexical	Is the signer performing a WH question word (e.g., WHO) without performing a question facial expression?	You performed a question word, but there was no QUESTION facial expression.	Use a "question" facial expression when performing a question word.	Face
error_ynq_beginning	Is the signer performing a YN-question facial expression that is not correctly aligned with the clause beginning?	YES/NO QUESTION facial expressions should start at the beginning of the clause.	The "yes/no question" facial expression should stop at the beginning of the question.	Face
error_ynq_end	Is the signer performing a YN-question facial expression that is not correctly aligned with the clause end?	YES/NO QUESTION facial expressions should stop at the end of the clause.	The "yes/no question" facial expression should stop at the end of the question.	Face

(Continued)

Table IV. Continued

Code	Description	Onscreen message text displayed to student in NOTES and POPUP conditions	Onscreen message text displayed to student in the SIMPLE or PHOTO conditions (Phase 2)	Where the arrow should aim for pop-up messages
correct_cond_if	Is a conditional facial expression happening now correctly during a conditional?	You have correctly used a TIME/IF facial expression at the beginning of a clause.	You have correctly used a "time/if" facial expression	Face
correct_ cond_time	Is a conditional facial expression happening now correctly during a time-phrase?	You have correctly used a TIME/IF facial expression during a time phrase at the beginning of a clause.	You have correctly used a "time/if" facial expression	Face
correct_ contrastive	Is contrastive role shifting (torso swivel) happening now correctly?	You appeared to use good contrastive structure (torso movement) during your signing.	You used body movement correctly.	Torso
correct_neg	Is a negative facial expression happening now correctly?	You have correctly used a NEGATIVE facial expression during a sentence.	You used the "negation" facial expression correctly.	Face
correct_point	Is the signer looking in the direction of pointing when establishing a reference point in space?	You correctly used eyegaze when pointing at a location in space.	You correctly looked at the location where you were pointing.	Face
correct_rhq	Is a rhetorical question facial expression being performed correctly during a question?	You appeared to use a RHETORICAL QUESTION facial expression during a clause.	You used the "rhetorical" facial expression correctly during a qustion.	Face
correct_topic	Is a Topic facial expression being performed correctly at the beginning of a clause?	You appeared to use a TOPIC facial expression correctly at the beginning of a clause.	You used the "topic" facial expression correctly at the beginning of a statement.	Face
correct_whq	Is a WH-question facial expression being performed correctly simultaneous to the occurrence of a WH-question word?	You have correctly used a WH-QUESTION facial expression during a sentence.	You have correctly used a "question" facial expression.	Face

ELECTRONIC APPENDIX

The electronic appendix for this article can be accessed in the ACM Digital Library.

ACKNOWLEDGMENTS

We are grateful to our research collaborators on this NSF-funded project, Professor Yingli Tian and her students at CUNY City College, who have provided video recording software for this study. We are also grateful to Sarah Ebling, Abhishek Mhatre, and Utsav Shah, who assisted in conducting some recording sessions with ASL students.

REFERENCES

Ben Bahan. 2016. The wolf who cried sheep. Online video. Retrieved from https://youtu.be/7Y44OUbwthQ.

Helen Cooper, Brian Holt, and Richard Bowden. 2011. Sign language recognition. In *Visual Analysis of Humans*. Springer, 539–562.

Gallaudet Research Institute. 2011. *Regional and National Summary Report of Data from the 2009–10 Annual Survey of Deaf and Hard of Hearing Children and Youth*.

David Goldberg, Dennis Looney, and Natalia Lusin. 2015. *Enrollments in Languages other than English in US Institutions of Higher Education*, Fall 2013. Modern Language Association.

Susan Goldin-Meadow and Rachel I. Mayberry. 2001. How do profoundly deaf children learn to read? *Learn Disabil Pract Res* 16, 4, 222–229.

Foad Hamidi and Melanie Baljko. 2013. Automatic speech recognition: A shifted role in early speech intervention? In *Proceedings of the SLPAT Workshop 2013*. 55–61.

Valerie Henderson, Seungyon Lee, Helene Brashear, Harley Hamilton, Thad Starner, and Steve Hamilton. 2005. Development of an ASL game for deaf children.In *Proceedings of the 2005 Conference on Interaction Design and Children* (Boulder, CO).

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 (ASSETS'15)*. ACM, New York, 139–146. DOI=http://dx.doi.org/10.1145/2700648.2809859

Pengfei Lu and Matt Huenerfauth. 2014. Collecting and evaluating the CUNY ASL corpus for research on American Sign Language animation. *Comput. Speech Lang.* 28, 3 (May 2014). 812–831. Elsevier. doi:10.1016/j.csl.2013.10.004

Dawn MacLaughlin, Carol Neidle, and David Greenfield. 2000. *SignStreamTM User's Guide, Version 2.0. American Sign Language Linguistic Research Project*, Report Number 9, Boston University, Boston, MA. Retrieved from http://www.bu.edu/asllrp/reports.html#RPT9.

Mark Marschark, Patricia Sapere, Carol M. Convertino, Connie Mayer, Loes Wauters, and Thomastine Sarchet. 2009. Are deaf students' reading challenges really about reading? *Am. Ann. Deaf* 154, 4, 357–370.

Rachel I. Mayberry and Ellen B. Eichen. 1991. The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *J Mem. Lang.* 30:486–498.

Ross E. Mitchell and Michael A. Karchmer. 2004. Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Lang Studies* 4, 2, 138–163.

Ross E. Mitchell, Travas A. Young, Bellamie Bachleda, and Michael A. Karchmer. 2006. How many people use ASL in the United States? Why estimates need updating. *Sign Lang Studies* 6, 3, 306–335.

Caio D. D. Monteiro, Ricardo Gutierrez-Osuna, and Frank M. Shipman. 2012. Design and evaluation of classifier for identifying sign language videos in video sharing sites. In *Proceedings of ASSETS'12*. 191–198. DOI=10.1145/2384916.2384950

Frank R. Lin, John K. Niparko, and Luigi Ferrucci. 2011. Hearing loss prevalence in the US. *Arch. Intern. Med.* 17, 1, 20, 1851–1852.

Carol Neidle, Judy Kegl, Dawn MacLaughlin, Ben Bahan, and Robert G. Lee. 2000. *The Syntax of ASL: Functional Categories and Hierarchical Structure.* Cambridge: MIT Press.

Russel S. Rosen. 2004. Beginning L2 production errors in ASL lexical phonology: A cognitive phonology model. *Sign Language & Linguistics* 7, 1, 31–61.

Jenny L. Singleton and Elissa L. Newport. 2004. When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognit. Psych.* 49, 4 (Dec. 2004), 370–407.

Patricia Spencer and R. Lederberg. 1997. Different modes, different models: Communication and language of young deaf children and their mothers. In *Communication and Language: Discoveries from Atypical Development*. M. Romski (Ed.). Harvard University Press. 203–230.

Michael Strong and Phillip M. Prinz. 1997. A study of the relationship between American Sign Language and English literacy. *J. Deaf Stud. Deaf Educ.* 2, 1, 37–46.

Carol Bloomquist Traxler. 2000. The Stanford achievement test, 9th edition: National norming and performance standards for deaf and hard-of-hearing students. *J. Deaf Stud. Deaf Educ.* 5, 4, 337–348.

Clayton Valli, Ceil Lucas, Kristin J. Mulrooney, and Miako Villanueva. 2011. *Linguistics of American Sign Language: An Introduction.* Gallaudet University Press.

Haijing Wang, Alexandra Stefan, S. Moradi, Vassilis Athitsos, Carol Neidle, and F. Kamangar. 2010. A system for large vocabulary sign search. In *Proceedings of the Workshop on Sign, Gesture and Activity*.

Kimberly A. Weaver, Thad Starner, and Harley Hamilton. 2010. An evaluation of video intelligibility for novice ASL learners on a mobile device. In *Proceedings of ASSETS'10*. 107–114. DOI=10.1145/1878803.1878824

Kimberly A. Weaver and Thad Starner. 2011. We need to communicate: Helping hearing parents of deaf children learn American Sign Language. In *Proceedings of ASSETS'11*. 91–98. DOI=10.1145/2049536.2049554

Ronnie B. Wilbur. 2000. The use of ASL to support the development of English and literacy. *J. Deaf Stud. Deaf Educ.* 5, 1, 81–104.

Received April 2016; accepted December 2016