

ASL Wiki: An Exploratory Interface for Crowdsourcing ASL Translations

Abraham Glasser
atg2036@rit.edu
Computing and Information Sciences
Rochester Institute of Technology
Rochester, New York, USA

Fyodor Minakov
fyodorominakov@gmail.com
Microsoft Research
Cambridge, MA, USA

Danielle Bragg
Danielle.Bragg@microsoft.com
Microsoft Research
Cambridge, MA, USA

ABSTRACT

The Deaf and Hard-of-hearing (DHH) community faces a lack of information in American Sign Language (ASL) and other signed languages. Most informational resources are text-based (e.g. books, encyclopedias, newspapers, magazines, etc.). Because DHH signers typically prefer ASL and are often less fluent in written English, text is often insufficient. At the same time, there is also a lack of large continuous sign language datasets from representative signers, which are essential to advancing sign language research and technology. In this work, we explore the possibility of crowdsourcing English-to-ASL translations to help address these barriers. To do this, we present a novel bilingual interface that enables the community to both contribute and consume translations. To shed light on the user experience with such an interface, we present a user study with 19 participants using the interface to both generate and consume content. To better understand the potential impact of the interface on translation quality, we also present a preliminary translation quality analysis. Our results suggest that DHH community members find real-world value in the interface, that the quality of translations is comparable to those created with state-of-the-art setups, and shed light on future research avenues.

CCS CONCEPTS

• **Human-centered computing** → **Accessibility systems and tools**; **Accessibility technologies**; **Empirical studies in collaborative and social computing**; • **Information systems** → **Collaborative and social computing systems and tools**; **Web interfaces**; *Web searching and information discovery*; • **Applied computing** → **Digital libraries and archives**; *E-learning*.

KEYWORDS

Deaf and Hard-of-Hearing, Sign Language, Bilingual, Interface, Education, Corpus, Crowdsourcing, ASL data collection

ACM Reference Format:

Abraham Glasser, Fyodor Minakov, and Danielle Bragg. 2022. ASL Wiki: An Exploratory Interface for Crowdsourcing ASL Translations. In *The 24th International ACM SIGACCESS Conference on Computers and Accessibility*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

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

© 2022 Association for Computing Machinery.

ACM ISBN 978-1-4503-9258-7/22/10...\$15.00

<https://doi.org/10.1145/3517428.3544827>

(ASSETS '22), October 23–26, 2022, Athens, Greece. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3517428.3544827>

1 INTRODUCTION

Approximately 1 in 6 adults in the U.S. is Deaf or Hard-of-Hearing (DHH), and prior literacy research shows that over 17% of deaf adults have "low literacy" [1]. Signed languages are the primary languages of Deaf communities worldwide, and they are completely distinct from local spoken/written languages. For example, American Sign Language (ASL) is the primary signed language used in the U.S., but it is a completely different language from English – not a one-to-one mapping. As a result, if a person is fluent in ASL, they are *not* necessarily fluent in English reading and writing. ASL is often DHH signers' primary language, and they typically prefer ASL over English, are more comfortable with, and understand content better in ASL [21]. Among this bilingual community, there is a wide range of literacy levels (e.g. studies have found fourth-grade reading levels among DHH high-school graduates [39]). Research has found that as a result there are lower educational outcomes among DHH individuals and lower rates of employment (and salaries) among DHH adults [2].

A major obstacle facing DHH signers is a lack of educational resources in sign language. Many educational resources are available in text (e.g. textbooks, literature books, online encyclopedias, etc.), but not in a signed language. As there is no standardized written form of ASL and sign language is typically in video form, these text-based interfaces do not adequately support users who prefer a signed language. Because of this lack of ASL content, DHH users often have to look up individual English words on a separate website or interface (e.g. English-to-ASL dictionaries) [10], and re-read the English content they are trying to consume [3]. Even though individual words can be looked up when necessary, this is not efficient, does not help to understand English grammar, and may be insufficient for DHH signers trying to understand English text. It would be helpful if an ASL version of the target content was available – having the entire sentence/article signed might be preferred by a DHH ASL signer rather than looking up individual words and/or re-reading multiple times.

At the same time, advancing sign language research and technology is currently impeded by lack of sign language data [8]. Existing ASL datasets typically offer a set of individual ASL signs, with their respective English meanings, and/or ASL glosses. They do not have sufficiently representative and diverse signers – they often consist of homogeneous sign language interpreters, small sets of signers, and poorly labeled videos of unverified quality (listed in [6]). In order to more fully understand and model the language, labelled continuous signing (i.e. complete sentences with annotations) from

diverse signers is needed. However, creating such a dataset is extremely difficult. It is not only expensive to produce, but it also requires a massive amount of human labeling and annotation, since there is no automated system to do so. It is also hard to enable a large pool of contributors, since most in-person data collection efforts are limited to individuals who live close-by within commuting distance, and have time in their schedules to contribute. How to enable everyday signers to efficiently contribute labelled continuous content, and how DHH users might respond to crowd-generated content remain open questions.

In this work, we present a novel interface that addresses two needs at once: 1) it provides a bilingual information resource and 2) it simultaneously generates a continuous labelled signing dataset that could be used by artificial intelligence researchers, ASL linguists, ASL learners, DHH ASL signers, and others. Our interface provides a side-by-side ASL (video) and English (text) synchronized interface, where users are able to read/view articles simultaneously in both languages. Users can also use this platform to contribute ASL translations of existing English texts in the communal database. For this exploratory work, we seed the interface with popular English Wikipedia articles, which are translated into ASL, and refer to this prototype system as "ASL Wiki". However, the same interface could be seeded with any long text, and could be used with any pair of written and signed languages. In terms of dataset creation, by enabling contributors to record segments of English text with known contents, the interface eliminates the need for humans to later segment and align the text and video. Such intensive labelling work is commonly done in creating parallel corpora containing signed language and spoken/written text (e.g. [17]).

To help understand the effectiveness of such an interface, we ran two exploratory studies. First, to better understand the user experience with the interface, we conducted an exploratory user study where 19 participants used the interface to consume and generate content, and shared feedback. Our results suggest that DHH individuals find real-world value in our interface, thought it was easy and intuitive to use, and were excited to see further development and identified several target audiences they would recommend the site to. Second, we also conducted an exploration into the quality of translations that can be generated through our interface. Results suggest that the translation quality is comparable to the quality of translations created through state-of-the-art setups for sign language translation. We conclude by discussing future work that this initial exploration introduces.

2 RELATED WORK

In this section, we focus on work relevant to our two motivations: supporting bilingual content, and supporting sign language data collection efforts.

2.1 Sign Language Educational Resources

Existing resources that make information available in a signed language compromise a small number of dictionaries ([11], [18]), educational materials ([14], [12], [20], [42]), lexical databases ([37], [19]), and mobile vocabulary apps ([27], [28]). Several examples of these are listed in [6]. The landscape of existing sign language resources is very small compared to the resources available for

spoken and written language users, who are typically considered by default. There have been limited attempts to create browser tools that provide signed translations of written content, to create signing avatars, and to more generally create recognition and translation systems [6]. These tools and resources are not viable due to the very limited amount of labeled signing videos with diverse, well-representative signers. There are some DHH content creators that strive to support accessibility of information and support the DHH community, such as the Daily Moth – a group who “deliver news in video using American Sign Language” [29]. However, as these efforts are sponsored and often composed of a small group of people, they are limited in the amount of content they can create, and often have to selectively offer a handful of content options – for instance, the Daily Moth says “the deaf host, Alex Abenchuchan, covers trending news stories and deaf topics on new shows Monday-Fridays”. Many people in the DHH community praise the Daily Moth due to the level of access it provides, being a bilingual information resource for selected news happening around the world [33, 41].

Our interface would enable crowdsourcing to address the problems of large-scale sign language data collection and diversity, naturalness, all while serving as a bilingual educational resource.

2.2 ASL and English Bilingualism

Prior work suggests that bilingual resources are useful for DHH fluent signers, rather than having any negative information-overload effects. Psychology researchers have established that it is not costly to switch from single to dual lexical retrieval (using two languages at once), and revealed a significant cost to turning off a language, which bilingual DHH users might do while trying to understand English text alone [16]. This suggests that an ASL and English bilingual interface, such as the one we have developed in this work, could be beneficial to DHH fluent signers by providing greater accessibility than English text alone.

The value of bilingualism in ASL and English has been further substantiated by Deaf-led organizations. The National Association of the Deaf (NAD), a nonprofit organization whose mission “is to preserve, protect and promote the civil, human and linguistic rights of deaf and hard of hearing people in the United States of America.” Internationally, NAD represents the U.S. to the World Federation of the Deaf (WFD), an international human rights organization. NAD supports bilingualism, using ASL and English, in the home and educational environment for DHH individuals. They advocate that bilingualism is important and effective because it fosters “positive self-esteem, confidence, resilience, and identity, factors necessary for lifelong learning and success” [30].

Despite the value of bilingual ASL/English resources, few exist. The Deaf Studies Digital Journal (DSDJ) “is the world’s first peer-reviewed journal dedicated to advancing the cultural, creative, and critical output of published work in and about sign languages and Deaf culture” [22]. It is a bilingual and bimodal publication primarily presented in both ASL and English. It features academic work in other sign languages, and offers scholarly articles, commentary, literature, visual arts, film/video, interviews, reviews, and archival history footage and commentary. To date, there have been 5 issues (spanning 2009-2020) with about 150 articles total. In the most

recent issue, each article has a split side-by-side view showing ASL (or other sign language) on the left, and English text on the right. The content is synchronized so that the English sentence being signed is highlighted. The video has controls so that the user can control playback of the signed video. Our interface builds on this, similarly providing side-by-side views in both languages.

To the best of our knowledge, there has been only one past attempt to systematically provide sign language translations of existing text. Signly¹ is a recent commercial effort to add “synchronous, in-vision, sign language translations on any webpage for any deaf sign language user”. They enable website visitors to select English text they would like translated into British Sign Language (BSL), which is sent to a professional interpreter for translation. Once the translation video is created, website visitors can click on the English text to trigger a pop-up translation video at the bottom-right corner. While this company helps make English texts online accessible, users have to request translations, and website creators have to contact and pay Signly to incorporate and maintain their services. Scale is also limited, as the translations are done by the Signly team. Our interface is similarly motivated to provide access to English text online. However, we enable crowdsourcing translations to streamline and scale data collection, and to enable a more diverse and representative group to contribute. We also display the text and video side-by-side in a more bilingual manner.

2.3 Sign Language Technologies Need Data

A recent paper [8] summarized the state of sign language processing. They hosted an interdisciplinary workshop with 39 domain experts with diverse backgrounds, where they reviewed the state-of-the-art, and listed calls to action for the research community. These calls included but were not limited to focusing on real-world applications and creating larger, more representative, public video datasets. They emphasized the current lack of data, cited as the biggest obstacle in sign language technology research. Data collection is difficult and costly, yet “without sufficient data, system performance will be limited and unlikely to meet the Deaf community’s standards”.

Despite these challenges, groups have worked on sign language data creation and curation. Datasets exist for many signed languages, including but not limited to German², American³, Argentinean⁴, Turkish⁵ (more listed in [8, 13, 31, 32], etc.). The main parameters of sign language datasets include the number of subjects, samples, language level, type, and annotations/labels. As explained in [8], existing sign language datasets greatly limit the robustness of systems trained on them. Current datasets are not sufficiently large – typically containing fewer than 100,000 articulated signs.

Also, many existing datasets contain individual signs, which may not be as useful for real-world use cases of sign language processing. For real-world applications, there needs to be natural conversation with complete sentences, i.e. “continuous” sign language. Continuous sign language recognition and translation are challenging due to epenthesis effects (insertion of extra features into signs) and co-articulation (ending of a sign affecting the start

of the next), among other difficulties. Solving these challenges requires large amounts of continuous sign language data to learn from. Continuous signed sentences would also be useful for DHH individuals trying to understand content, especially new concepts, as it is natural and comfortable for them. There are some continuous signing datasets, such as [15], which help fill this void. However, these datasets are typically small and recorded in a laboratory or studio environment, rather than a natural setting, which makes generalizing to diverse users and real-world environments difficult.

Currently, the process of producing large ASL datasets is prohibitively expensive, due to the equipment needed and the time it takes to collect and label/annotate ASL data. Crowdsourcing is a more affordable alternative to traditional in-lab collection, and has been successfully used in other accessibility domains. For example, research has explored collecting images/videos and questions from blind and low-vision users, with answers provided by the crowd [4, 5]. Other accessibility crowdsourcing research has explored providing image alt-text [35, 36], transportation information [34, 38], and live captions [23, 24]. Our current work adds to this body of work, by providing an initial exploration into crowdsourcing continuous sign language data.

Along with the lack of bilingual information resources, lack of data motivates our “ASL Wiki” interface.

3 ASL WIKI PROTOTYPE

We have created “ASL Wiki” – a prototype site where people can crowdsource ASL translations of English articles, providing a community resource that supports accessibility as a bilingual information resource, while also tackling the lack of continuous ASL datasets with English labels. In this section, we describe our prototype and design process.

3.1 Design Process and Criteria

We engaged in an iterative design process to arrive at our “ASL Wiki” website design. We first identified design criteria the platform needed to meet (e.g. that the text used is available for use on the platform and in a dataset, that participants can contribute remotely without specialized hardware, and that translations are segmented and labelled). With these identified, we started with drawn designs which were iteratively refined and implemented. Throughout the process, we continued to meet with stakeholders consisting of a group of interdisciplinary Deaf and hearing individuals who have deep ties with the DHH community and incorporate their input. These stakeholders tried out the evolving prototype, and also discussed the project and provided guidance.

Through our meetings, we chose to explore creating and reading bilingual versions of Wikipedia articles, rather than play scripts, books, or other resources. We decided on Wikipedia articles because they are generally neutral, publicly available, and popular informational resources. There also exist other parallel corpora of Wikipedia content which have been useful for natural language processing and artificial intelligence/machine learning.

Our iterative design process uncovered specific user requirements of our interface. We found that the interface needed to show ASL and English at the same time, so that users could see both and easily look at one or the other as they wished. Our interface

¹<https://signly.co/>

²<https://www.phonetik.uni-muenchen.de/forschung/Bas/SIGNUM/>

³<https://github.com/YAYAYru/sign-language-datasets>

⁴<http://sedici.unlp.edu.ar/handle/10915/56764>

⁵https://www.cmpe.boun.edu.tr/pilab/BosphorusSign/home_en.html

also needed to show which English portion is being signed in the current ASL video, so that users can keep track of their position in both the video timeline and the English article. Users who are recording their videos should have an efficient, streamlined way to record sentences, meaning that the interface should not pose unnecessary overhead. It should be allowed for multiple people to submit recordings for the same English sentence, as different people might sign differently (e.g. regional accents or varied interpretations), or have preferred signs for specific English words.

3.2 "ASL Wiki" Design

3.2.1 Homepage. We took inspiration from the idea that Wikipedia is a "free content, multilingual online encyclopedia written and maintained by a community of volunteers" [40]. On the homepage of the ASL Wiki site, on the left hand side, is a checkbox list of featured categories. Users can use these checkboxes to bring up relevant articles, which appear in the middle with fractions indicating how many sentences there are in the article, and how many of these sentences have been recorded by at least one user. Being clickable, the rows of article titles also display a "Record" button that takes you to the reading/recording interface (discussed further in subsequent section).

On the top of the homepage is an introductory title and paragraph, along with an ASL video of someone signing this text. Once you are logged in, on the top right of the page is a button that allows you to view and edit your profile, or sign out. Next to this button is a gamifying trophy icon displaying the number of sentences the logged-in user has recorded. This was added as it is a common element of social media sites that display the number of "posts" an user has submitted. It potentially incentivizes the user by showing them how much sentences they have recorded.

In the middle of the page, between the top banner and article table of contents, is a numbered instruction summary to remind users how to navigate and use the interface. Especially as users are able to leave the site and come back later, and since they navigate into and out of specific articles, they may need a persistent reminder of how to use the site, which is why we added this. A screenshot of the homepage is shown in Figure 1. Once the user selects an article, they are taken to the reading/recording view. This view has a toggle on the top to switch between the recording view and the reading view.

3.2.2 Recording and Reading View. In both recording view and reading view, the main layout is the same: it is a split, side-by-side bilingual interface. On the left is a placeholder for an ASL video. On the right is the article in English.

If in record view, once the user selects a sentence on the article, the ASL video placeholder becomes a self-view of the user's webcam, so that they can see themselves. Their self-view is overlaid with a head and body guide to encourage users to center themselves in the recording. A 3-second countdown commences, and then the user would sign the English sentence in ASL. While they are recording, the according English sentence is highlighted, to mark and keep track of their place in the article. When they are finished, clicking on a stop button underneath their self-view stops the recording, and displays their recorded video for playback. If the user approves, clicking "Keep" will submit the video, and auto-progress to the next

sentence in the article, or clicking "Redo" will prompt them to redo the recording. The English sentences that have been recorded will show a video camera icon. There is also a guide on the top, above the English article, to remind users how to use the interface. Also, underneath the ASL video placeholder is a picture demonstrating a good recording setup and a bad example, to remind users that they should be sure to position their webcams so it captures their upper body and that their arms/hands do not go out of frame while signing. There is also an upvote/downvote button where users can give feedback on the ASL video.

In reading view, the site enables users to access parallel content in ASL and English. After toggling to reading view, the same English article is kept, and now shows a "play" icon next to the sentences that have been recorded. Clicking on a sentence will highlight that sentence, and play the respective ASL video. Once the video completes, it auto-progresses to the next sentence. There is a playback control underneath the video so that the user can go back, forward, redo, pause/play, and control the playback speed. There is also a toggle to turn on or off the auto-progression. It is possible that multiple users would sign the same English sentence, so underneath the ASL video is a list of the users who submitted videos for the currently selected English sentence. The user has the option to switch between signers if they desire. A screenshot of a sample reading view is shown in Figure 2.

4 USER STUDY

To explore the usability of our ASL Wiki site design, we ran a remote user study, with Institutional Review Board (IRB) approval. In this user study, participants answered survey questions, tried out the reading and recording views, and discussed interview questions about their experience.

4.1 Participants

4.1.1 Recruitment. Participants were recruited via mailing lists, social media posts, and snowball sampling. The recruitment criteria was that they use ASL, are 18 or above years of age, and have a computer with a webcam. 19 participants were recruited in total. The sessions ran for about 1 hour, and participants were given a \$30 (USD) Amazon gift card for their participation.

4.1.2 Demographics. Out of the 19 participants, 15 identified as Deaf, 3 deaf, and 1 Hard-of-Hearing. 11 identified as female, and 8 male. The average age of all participants was 26.1 with standard deviation 2.2.

Participants self-reported their ASL fluency on a scale from 1 (I do not use ASL) to 7 (I am fluent). The average fluency was 6.4 with a standard deviation of 1. Generally, all participants were educated, with only 3 out of 19 not having a bachelor's degree yet at the time of participation. Participants were diverse, with 12 self-identifying as White (e.g. German, Irish, English, Italian, Polish, French, etc), 5 as Asian (e.g. Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, etc.), 1 as Black or African American (e.g. African American, Jamaican, Haitian, Nigerian, Ethiopian, Somalian, etc), and 1 as Middle Eastern or North African (e.g. Lebanese, Iranian, Egyptian, Syrian, Moroccan, Algerian, etc). The 19 participants came from 8 different U.S. states.

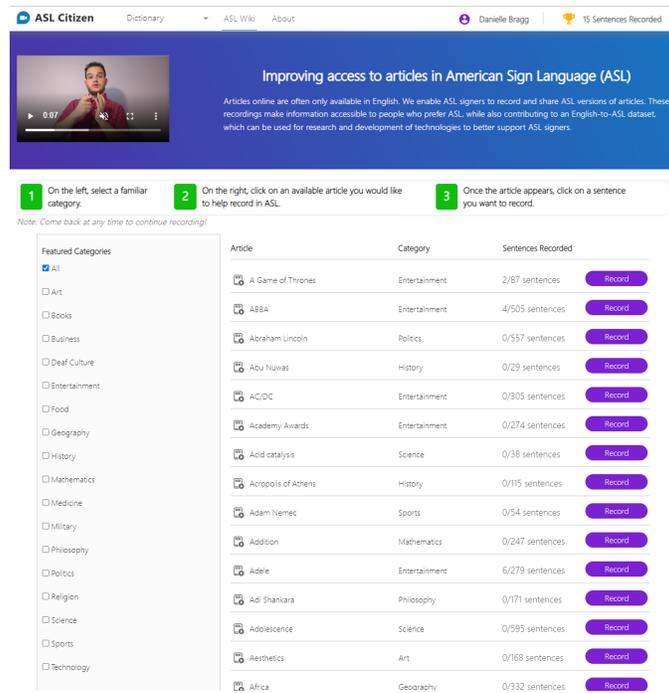


Figure 1: Screenshot of ASL Wiki homepage.

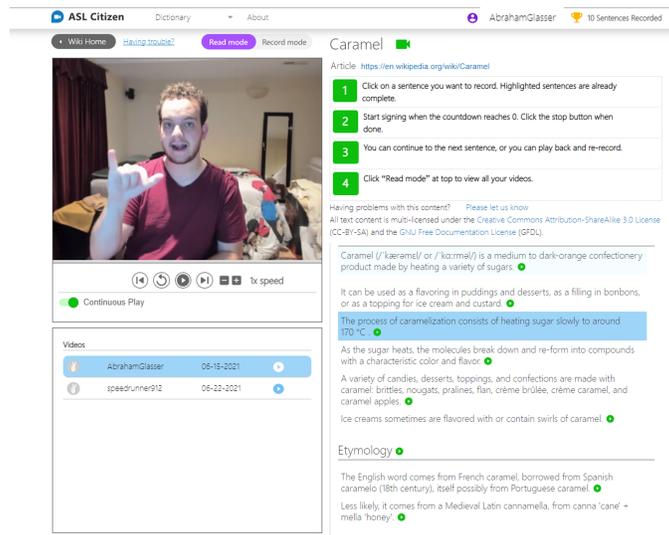


Figure 2: Screenshot of reading view of article "Caramel".

4.1.3 *Prior Experience with ASL and English.* All participants reported that they read English text online daily (n=5) or multiple times a day (n=14). It was reported in the demographics survey that participants read English text via websites, books, articles, video transcripts, and social media posts. Along with these 5 options, we had also listed podcasts (and "other") as the answer-choices on this survey question, but nobody selected that.

Participants were asked the question "How often have you encountered websites you wish provided ASL videos instead of or in addition to English text?" 3 answered "multiple times a day", 5 "daily", 5 "weekly", 2 "monthly", 2 "less than once a year", and 2 "never".

All participants except one said that they watch ASL videos online frequently (1 said "yearly", 3 "monthly", 5 "weekly", 5 "daily", and 5 "multiple times a day"), typically through video blog (vlog)

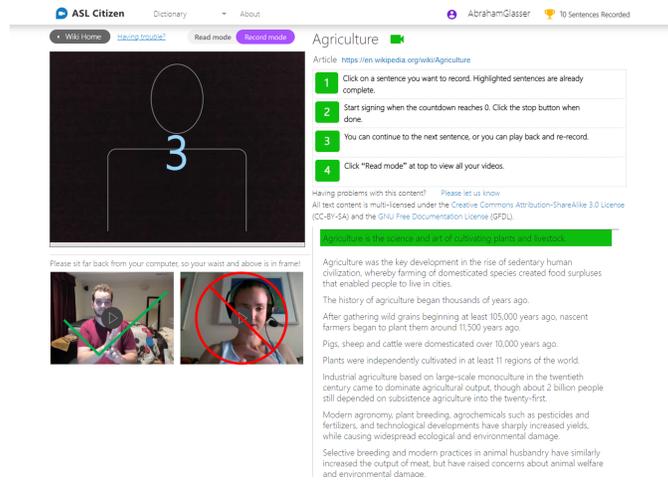


Figure 3: Screenshot of recording view of article "Agriculture".

posts, YouTube videos and other social media videos. Participants commented that they have seen content on various social media platforms where someone is signing in ASL, and there are English captions visible, so they have seen bilingual/bimodal content before, and are comfortable with it. 10 out of 19 participants said that they have at least once created content like this that had both ASL and English, and that they created the ASL video first and added English subtitles afterwards. 9 out of these 10 did this to post on social media, where they have both DHH and hearing friends, and 1 said they only did it for a homework assignment or class project in college.

4.2 Procedure

An online form walked participants through the study procedures while a DHH fluent ASL signer was on a video call with the participant. Each participant was scheduled for their own session, and the entire procedures took approximately 1 hour. The procedures were as follows:

- (1) **Consent:** Participants engaged in a consent process with IRB-approved language through the online form. The researcher on the video call checked whether the participant needed any portion of the consent language signed in ASL so that it was fully understood.
- (2) **Background:** Through the online form, participants were asked multiple-choice questions about their prior experience with using English and ASL online.
- (3) **Reading:** After this, they followed instructions on how to access the ASL Wiki site and sign in, and were directed to the "Caramel" article which had been entirely pre-recorded by a DHH ASL signer from our research team. They engaged with the interface to read this article until they were satisfied.
- (4) **Recording:** Next, they were instructed to select any article of their choice and record themselves signing. Since we wanted to closely match a real-world experience of using our site, participants were given the flexibility to record as much (or as little) as they wanted to, but were told to use the recording interface until they were confident that they

got and understood the full experience of recording and contributing to the site, and were told they would discuss their experience afterwards.

- (5) **Semi-structured interview:** While the fluent DHH ASL signer continued to be on a video call with them, they engaged in a semi-structured interview with guiding questions spanning short answer, long answer, and Likert-scaled question items. The interview focused first on the reading view, asking about their experience and understandability using the interface, and then were asked questions about their experience and challenges (if any) while recording. Lastly, questions were asked about the overall concept of the site, what they liked and disliked, and whether they would recommend the site to others. Appendix A.1 provides our interview questions.
- (6) **Demographics:** After the interview portion, participants returned to the online form where they filled out demographics and compensation information.

5 USER STUDY RESULTS

We discussed with each of the 19 DHH participants during the experiment to gauge their reactions and experiences with the reading and recording views of the interface. We evaluated how they used the site, to understand their motivations, challenges and strategies, and the benefits they took away from the site. We thematically analyzed the interview responses and performed statistical analysis of their responses to the questionnaire. We also collected feedback and identified several target audiences who the users would recommend the interface to.

5.1 Reading View

5.1.1 ASL vs. English. Participants valued having both English and ASL versions of articles available for consumption. On average, participants self-reported that, while reading the "Caramel" article that our research team had entirely pre-recorded, they looked at the English part 65% of the time, and the ASL video 35% of the

time. Participants explained that the English part is faster to read, with P1 saying "...it's faster to read and skim through. It's more of a habit because I'm used to reading English articles". P8, who reported looking at the English part 50% of the time, said "I like ASL. [It] is more visual and I can visualize it better, but for English I can read it faster. If I just want to consume the content and save time, I would look at the English 100%. If I wanted to fully understand, learn, visualize, maybe 50/50 – I'd also be curious what it looks like in ASL".

Participants were asked to indicate how understandable the ASL content and the English content in the "Caramel" article they viewed was, on a scale from 1 (very difficult) to 5. For the ASL content, the average was 4.6 (s.d. .7), and 4.8 (s.d. .4) for the English part. Even though participants said that the English part was very easy to understand, all 19 participants answered "Yes" to the question "Was it helpful to view the content in both English and ASL?" P11 explained "Yes, I can imagine how it would be helpful for the general. It's a nice tool for me to use, and I would like having it even if I don't use it much".

5.1.2 Interface Usability. On a scale from 1 (very difficult) to 5 (very easy), participants said that the interface was very easy to use, giving it an average of 4.5. Most of the difficulty came from not having prior experience and not knowing what to expect with the interface, e.g. P10 saying "I didn't think there was any information overload – in the beginning I wasn't fully sure what to do. Maybe the first sentence could be highlighted with the video, that would make it more clear there is ASL there". P2 commented "at first when I opened it, I wasn't sure what to do – my eyes caught the English part first, and I ignored the left half – and then it took me a while to realize that the left side was empty until I clicked on some text, and the video player showed up. [...] I think there should have been some kind of tutor/illustrations with directions of how to use this site before I went ahead and looked at an actual article".

Most participants (12 out of 19) did not use the upvote/downvote button that was available to them while viewing the "Caramel" article. Some participants said that they did not see it, while some did but decided not to use it. P12 said that they do not use it in general, such as upvoting/downvoting on Reddit, liking/disliking on Facebook or YouTube. P9 said "I didn't know about the feature until I arrived at this question. I normally skim through contents", and P14 said "I wasn't focusing on providing feedback on performance". Those who did use it generally said that they wanted to give feedback, with P4 saying "I wanted to give feedback on the video, so I clicked yes – I noticed the signing was clear and matched the English so I went ahead and clicked yes". Some participants such as P6 emphasized it was important: "I think it's important to use, yes I would use it, it gives feedback to other people and I can help this website advance and develop in the future and make sure it has good content", and P16 suggested it would prevent misunderstandings, saying "I don't want some signers to use wrong signs or say it in the wrong way which will make viewers misunderstand. We want to avoid that", and P11 made an analogy to real-world applications they've seen, reminiscing "Yes, it's the same as FAQs or other articles that say "was this article helpful?" – this is the same situation".

We collected some feedback about the interface, to understand how our interface could be improved and help inform future work on such bilingual interfaces. These feedbacks typically consisted

of user interface preferences and suggestions, such as coloring and layout styling. There were also some suggestions about the fundamental system. P7 suggested a different layout: "For me, I would prefer top and bottom rather than side by side, so it's kind of like captions. It was a little challenging for me to have it side by side". P13 suggested making the recorded videos easier to find: "One suggestion I have is that it might be nice to have a separate scrolling bar other than the browser one where it'll indicate the recorded statement bits. E.g. code changes in a code review". Besides these feedbacks, users also complimented the interface, P4 said "I liked the clarity, green highlight, follow each other, I liked the time/playback, matching", P13 "What I liked about the interface is that each statement and section is reasonably spaced out which makes it easier to read and I like how there's a clear indicator whether if there's recordings for it or not".

5.2 Recording View

A total of 202 sentences were recorded from our 19 participants. On average, participants recorded 11 sentences. Participants recorded in 25 total articles from the Entertainment, Deaf Culture, Sports, Books, Mathematics, Technology, Food, Geography, Art, and Politics categories.

5.2.1 Challenges and Strategies. Participants were asked if they found any content challenging to record. They reported that they generally selected articles from topics they thought they were the most familiar and comfortable with. For instance, P11 said "I picked the content I was most comfortable with, and it was straightforward and just facts, so it wasn't challenging. I can imagine if I picked a STEM article or something complicated it would be challenging". Some participants commented that they felt it was challenging to actually translate the English content into ASL, because they were not sure how to sign some words, or were not sure how to make it so it wasn't a word-for-word English-to-ASL translation, but rather a concept-to-concept translation – P18 said "it can be a bit challenging to keep it simple and brief yet informative", P16 reflected that there were "some words that I'm not sure if they have signs for them", and P3 summarized "sometimes I have to reread and think about how I will sign it to try not to be too English".

We asked participants if they had any strategies they employed while recording content. Most said they did not – they commented about trying not to be too "English" in their signing, with P1 saying "I would read first, and then think about my understanding of it, and try my best to explain it in ASL. I wanted to avoid one to one or exact English translations" and P7 commenting "I tried to find simpler sentences, but most sentences required a lot of fingerspelling. It was challenging to use it, I didn't really think through it, I just read the English part a couple times and then tried my best". We noticed that not all participants started at the top of their selected articles. It seems that some participants selected sentences throughout their articles and did not always record consecutively.

5.2.2 Interface Usability. To ensure whether the interface itself caused any significant issues for participants trying to record themselves, we asked them to rate, on a scale from 1 (very difficult) to 5 (very easy), how easy the interface was to use. For the 19 participants, the average was 4.6. It did not appear that the interface caused any further challenges to the recording experience, with

P1 saying *"I thought it was straightforward and simple"*, P3 *"...liked Redo/keep, add playback/review to watch it before deciding"*. Some participants had suggestions about the interface to make their experience better, such as the seemingly abrupt countdown that started as soon as they clicked on an English sentence to record, but there were conflicting responses as some participants said they disliked it, e.g. P7 suggesting *"maybe instead of auto countdown, I felt more pressure, I would rather click on the sentence and then have a record button"* and P13 who said *"I dislike that I can't manually start recording"*, while some liked it, e.g. P11 *"... I liked the countdown, 3-2-1"*. Participants also made some suggestions for extra features, such as being able to trim the video before submitting, moving the placement of the self-view, an explicit way to "skip" (rather than "redo" or leaving the page).

5.3 General Experience

5.3.1 Enthusiasm. After participants had tried out our bilingual interface, we asked whether they "wish more content online provided both English and ASL?" from a scale from 1 (strongly disagree) to 5 (strongly agree). The average response was 4.6 (s.d. .76). Participants gave examples of where they have wished they had access to both English and ASL. These examples included but were not limited to news, podcasts, articles, social media and entertainment. Participants mentioned the Daily Moth, where they have seen both ASL and English captions or transcripts, but they mentioned that these are selected specific news, and they wish they had access to a more broad, general selection of news around the world. Some participants mentioned that they wish they had this kind of bilingual resource when they were learning about things for their classes, projects, and homework. These findings suggest that people may want to use a tool, similar to our novel interface, in the real-world.

Despite the 19 participants' desire of having more content online provided in both English and ASL, they were not as interested in generating this sort of content themselves. When they were asked "Would you be interested in generating content available in both English and ASL? (1-5: Strongly disagree– Strongly agree)", their average response was 3.6 (s.d. 1.6). Some of their rationale included not wanting to record themselves and/or posting publicly, with P1 saying *"I personally would not, because I personally don't like recording myself and posting online publicly"*, P10 *"No, because I feel like I'm signing wrong, or feel that people would judge my signing for being English, etc"*, P17 *"No, I'm a camera shy"*. Participants who indicated that they are interested were inclined to do so because they felt they would be giving back to the community, and supporting this concept of accessibility, e.g. P2 *"Yes, I wouldn't mind – because I feel like there is a lot of ASL content out there that is not neutral, where the people who are signing are biased, or give biased information. This would be nice and I would like to help increase access while still keeping neutral and spreading information in a neutral manner"* and P3 *"Yes, because if I can get access like this, why not I give back, I don't want others to miss out"*.

5.3.2 Personal benefits. One participant mentioned that the site would benefit them because they can use it while teaching, to make sure their students have access and can understand the content fully. Many participants mentioned the site would help them understand content better and go through content faster, since they

wouldn't have to spend time looking up specific English words in a separate interface and/or re-reading the English text multiple times. Some participants said that this would also help them improve their signing and presentation skills, since they could benefit from watching their own videos, or pick up new signs for unfamiliar words. For instance, P14 said *"I can improve how best I can interpret English in ASL"*, P4 similarly saying *"If I record, I could benefit from watching my own videos, I will see if I signed it clear and understood it well. I would also benefit from reading myself, and others would benefit by reading my videos that I contribute"*. When we asked them what kind of content they would like to see on the website, they mentioned things they were studying, e.g. P2 *"related to my major, tutorials on 3d design software, art, technology, art terminology, for example gothic art history, etc."*, things they were interested in, as P11 brought up *"nutrition, diets, women's health, for example there's a lot of things that are related to hormones, specific foods affecting things, having ASL there would be nice"*, general news, information and topics, with P8 saying *"news, health, podcasts, could be a safe place for community involvement, like an area for people to post news around the world, gaming area, etc. Make subcommunities for gamers, etc, same concept as Facebook groups, Reddit subreddits, etc. But everyone is deaf and uses ASL"*, among several others. Many participants were very supportive of the idea, and did not care what kind of content is available, as long as a lot is available, with P19 saying *"... every site should have this option, all kind of topics are welcomed"*, P8 agreeing *"as much as possible, no limits"*, P5 *"there's so many topics to choose from, I would just pick the best and most informative articles for education"*, and P7 *"not that I can think of, general Wikipedia articles would be good"*. This shows that participants were very supportive of the site.

5.3.3 Concerns. We asked the participants whether they had any concerns while using a site like this. 7 out of 19 participants explicitly brought up the concern that there wasn't control over the quality of users' submitted videos. For instance, many commented that people may not have professional backgrounds, or that they may have something inappropriate or unintentional (such as other people) in the backgrounds of the videos they submit. People also mentioned that users have varying devices and webcam technologies, so that the quality of the videos themselves may not be as good as they'd like – perhaps the lighting would be bad, the video would be choppy or blurry, etc. A few participants also mentioned that the site may find users who misinterpret or inaccurately translate content. P18 said *"It can be misinterpreted easily if the translator is not professional or a novice"*, P11 brought up that *"not everyone can translate well, so that would be my concern – there might be some bad videos. I recommend having STEM topics assigned to people who are specialists in that field"*. A participant also brought up the issue of privacy, stating that they are concerned about the privacy of their data, and who would "own" it and who would be able to access it, especially if it was public.

5.3.4 Participant Impression. Overall, participants said that they enjoyed using the website, and that they thought it was "cool to use". When asked "Would you want to use a website like this to read content in the future?", 14 participants said "Yes", and 5 said "No". The participants who said "No" said that they are already comfortable with reading English text alone, and do not require

ASL for reading comprehension. Despite this, the participants, on average, said 4.5 to the question "How likely are you to recommend this website to others? (1-5: very unlikely – very likely)". Participants suggested many different groups of people who they would recommend this site to. They would recommend it to DHH individuals, because of the communication barriers they face, as P2 said *"I feel for DHH people, and others who are not good at English, and have communication barriers and have a lack of education, they can learn well through this site"*, P7 said *"I would recommend it to people who I know grew up signing and struggle with English, they could improve their reading skills and understand content better"*, P16 *"pretty much everyone with ASL especially for people who have weak English skills"*, and P4 bringing up *"international friends who don't know English very well, it would help understand English and ASL, or other people whose first language isn't English"*.

Many participants mentioned they would recommend the site to people who are learning ASL, since the site is bilingual and has synchronized English and ASL content, as P3 says *"friends who are interpreters, on their own time learn ASL/translation, receptive skills, signing skills"*, P15 *"ASL students for learning and people who are thriving to learn ASL"*, P17 *"If the website has enough recordings or gains popularly among users, I would recommend to a friend who isn't fluent in ASL"*, and P19 summarizing both ASL learners and the DHH community *"this would be great for people learning ASL. They can practice their receptive skills, and learn how to follow the ASL grammar structure and sign placements. This is also good for every person in the deaf community who may prefer reading captions only some days and ASL other days, or anyone who has a preference in how they absorb information"*. Since the site has been seeded with articles from Wikipedia, which are normally informative and neutral, participants suggested people who often look up information, or use information in their profession, such as researchers, with P6 saying *"school educational use, like for students to do research, or college students/professionals to record videos, k-12, community college, ..."*, P5 suggesting teachers *"I would recommend it to teachers. I think the website would be best for education and is very educational rather than recreational, so teachers could use it to record content and provide information online"*, and P8 recommending learners: *"Maybe people who want to learn more things, learners, people who typically look stuff up and read things"*.

6 TRANSLATION QUALITY EXPLORATION

While our ASL Wiki site was designed to facilitate translation contributions, how the interface design may impact translation quality is unclear. Interpreters typically generate ASL translations of English texts in large sections (e.g. paragraphs). In contrast, our interface elicits of text segmented into sentences to enable readers to access spot-translations within long texts. Our interface also provides built-in mechanisms to facilitate the translation process (e.g. marking completion progress within the text, providing the text and recording interface in the same tool).

To explore the potential impact of the interface on translation quality, we ran a small experiment comparing a set of recordings generated through our interface to a comparable set generated through state-of-the-art recording setups. Specifically, we paid four professional Deaf interpreters to record 20 articles in both setups,

and then paid two fluent ASL experts to evaluate all the recordings, and compared the results. Our results suggest that the quality of translations created through ASL Wiki are comparable to those created through standard state-of-the-art setups, with potential slight improvements to translation accuracy and recording quality.

6.1 Procedure

6.1.1 Video Generation. We paid four Certified Deaf Interpreters (CDIs)⁶ to translate a set of 20 Wikipedia articles twice – with both our interface and with their standard translation setup. We chose to work with professional Deaf interpreters in order to enable comparison to state-of-the-art translations. Each CDI was assigned 5 articles to record twice, and we counterbalanced the procedure, so that two CDIs started with our interface and then used their standard setup, and the other two CDIs did the reverse.

In the standard recording procedure, the interpreters were given access to the plain text, and asked to record a translation of the text in sections. They were instructed to use their typical setup and procedures for such jobs – for example, referencing the text and/or personal notes and recording through a video camera app on their laptop or smartphone. This is a standard type of translation job taken on by professional ASL interpreters (e.g. to translate written questions in a survey, or to translate consent form language).

Each CDI translated their own set of 5 Wikipedia articles. Each set spanned a variety of topics, including both technical and non-technical topics. In total, 17 topics were covered in these 20 articles (identified through topic modeling on the most popular 810 English Wikipedia articles): Geography, Entertainment, Sports, Deaf Culture, History, Science, Mathematics, Medicine, Business, Politics, Technology, Military, Philosophy, Food, Books, Religion, Art. Article length ranged from 105-627 words (avg 309), and from 4-29 sentences (avg 15). In total, we collected 308 recordings through our ASL Wiki interface (corresponding to individual sentences), and 111 recordings through state-of-the-art interpreter setups (corresponding to sections).

6.1.2 Video Evaluation. To compare the quality of the two recording sets, we paid two fluent ASL linguists to evaluate each video along five dimensions. These dimensions capture the accuracy of the translation from English to ASL (Q1), the quality of the ASL independent of the English (Q2-Q3), and the completeness of the data captured (Q4-Q5). The dimensions and exact questions that the experts answered for each video are listed below. In addition, the experts had the opportunity to enter additional notes for each video, and we also engaged in a debrief meeting to gather their feedback and observations about the video sets as a whole.

- Q1)** Translation accuracy: How well does the ASL recording convey the meaning of the English? (Scale of 1-5)
- Q2)** Linguistic correctness: How correct is the ASL execution (e.g. were there many mistakes with handshape, movement, grammar, etc.)? (Scale of 1-5)
- Q3)** Signing naturalness: How natural is the ASL (i.e. how similar is it to ASL you might run into in real life)? (Scale of 1-5)
- Q4)** Recording quality: How good is the recording quality (e.g. is it blurry, is the lighting good, etc.)? (Scale of 1-5)

⁶<https://rid.org/rid-certification-overview/available-certification/cdi-certification/>

Q5) Signing captured: Is the full signing space captured in the video (i.e. hands, torso, surrounding are)? (Yes/No)

6.2 Results

The expert evaluations of the recordings generated through our interface and through the CDIs' standard setups were comparable, across all five explored dimensions. Figure 4 shows the overall results – average score and standard error for Q1-Q4, and the percent of videos that were evaluated as having captured the full signing space for Q5. We ran two-sided Wilcoxon-Mann-Whitney tests with Bonferonni correction to compare evaluations of the two interfaces for Q1-4. For Q1 (Translation accuracy), Q2 (Linguistic correctness) and Q3 (Signing naturalness), there was no statistically significant difference ($p > .0125$). For Q4 (Recording quality), the test showed statistical significance Q4 ($U=83175.5$, $p < .005$). We also ran a χ^2 test to compare Q5-Signing captured, which was not statistically significant ($p > .05$).

During our debriefing, the expert evaluators identified some patterns in the data. They noted that the recordings, in particular those created through ASL Wiki, contained straight translations rather than interpretations. For example, the interpreters did not tend to elaborate on concepts from the text to ensure that the meaning in ASL is clear, or to provide additional context not provided in the text. Instead, they tended to stick to the exact text. They also noted some examples where it seemed that the interpreters had not done the full prep work to understand the content they were translating. For example, this was evident to one expert in a translation of some plant anatomy, which lacked the appropriate visual representation. One expert also noted that they had expected to see a larger difference in quality between the recording sets, in particular due to the difference in text segmentation lengths. They were surprised that there was not a larger difference in translation accuracy and quality for the longer and shorter excerpts.

7 DISCUSSION

Generally, the results of our exploratory studies suggest that it may be possible to use specialized interfaces to crowdsource ASL translations of English text, to provide valuable bilingual resources to the community and to curate ASL data. Our user study results suggest that users would find value in such a bilingual ASL and English platform, and would be willing to contribute, especially if incentivized. At the same time, our translation quality exploration suggests that the interface enables high-quality translations. In this section, we provide further discussion on our exploratory work, the limitations of this initial work in this space, and related future work.

7.1 User Experience

Because participants were not incentivized further for contributing more videos during our user study, the majority of participants only contributed until they figured out and were satisfied with the user interface and experience for the recording view, with an average of 11 sentences per user. It seemed that participants generally chose to contribute to topics that were personally meaningful to them, especially those who contributed a larger number of recordings. It is possible that an expanded range of topics that interest more people

would thus incentivize contributions from the community. Further incentivization, such as credit for class or monetary payment could also be beneficial to deployment at scale.

Participants indicated that the reading and recording interfaces of our website design were easy to use. Even though participants all thought the site was easy and intuitive to use, several would rather only use it for reading bilingual content, rather than contributing ASL videos. They thought it was helpful to view content in both English and ASL, and mentioned several cases where they wish they had this level of accessibility in media. They talked about some of the challenges and strategies used while recording. The website was strongly supported and all participants identified populations that they would recommend the site to. Participants also suggested many different topics that could be added to the interface that would benefit them and others in mind.

Even participants who commented that they were fluent in English and ASL still indicated that seeing content in a bilingual, bimodal form was useful. Even if it did not help them understand the content itself better, some participants still mentioned that they could pick up new signs or improve their signing and presentation skills. Overall, participants enjoyed using the website, and identified several use cases and target audiences who they would highly recommend the interface to.

During the interview portion of the user study, we collected feedback from participants so that we could further iterate upon our design. These feedbacks would also be useful for future researchers who want to generalize our interface, and potentially use it for other signed or written languages. While our exploratory user study serves as a proof-of-concept, several research questions have arisen. We have identified several research avenues and next steps as a result of this work.

7.2 Translation Quality

In our translation quality exploration, it is possible that linguistic correctness was slightly more reliable with our interface due to reduced cognitive load. Our interface provided required shorter excerpts of text to be translated. It also simplified the recording task by keeping track of where the user was within the text, auto-progressing to the next excerpt, and providing the text and video feedback in a single interface rather than requiring the interpreter to manage two separate interfaces for these components. It is also possible that the recording quality was slightly better on average with our interface because the quality of the recording was less dependent on the quality of apps that the interpreter has available to them. While we did not provide hardware, we provided built-in recording software in our website, unlike state-of-the-art setups that are dependent on the recording software that interpreters have access to and know how to use.

While our exploration suggests comparable translation quality with ASL Wiki compared to state-of-the-art translation setups, it still leaves open questions about the impact of isolated interface components. For example, it would be interesting to examine the effect of different text segmentations within our interface, possibly ranging from individual words, to sentences, to paragraphs or sections. Similarly, it would be interesting to experiment with the effect of different types of visual cues for orienting the translator within

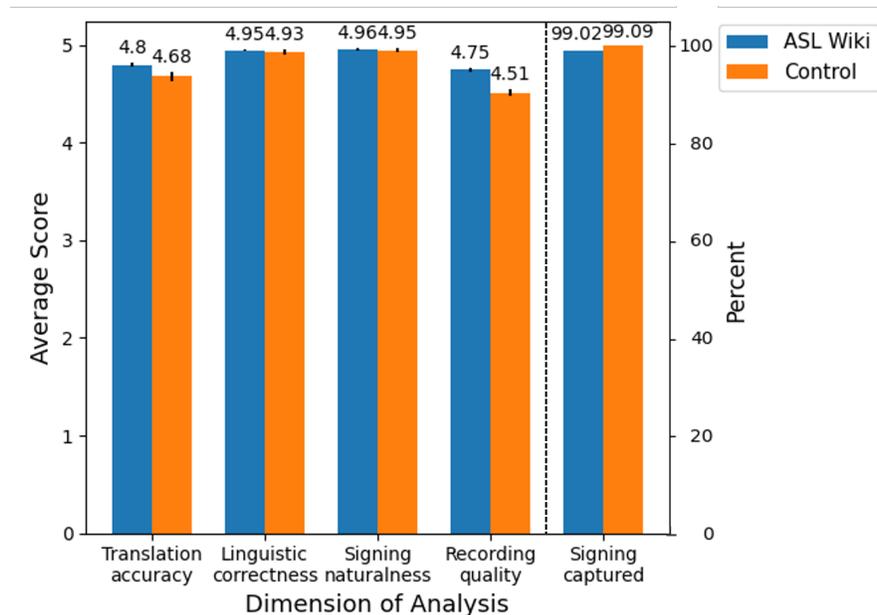


Figure 4: Comparison of expert evaluations of ASL translations of 20 Wikipedia articles, recorded by CDIs through ASL Wiki and a control state-of-the-art setup. For Q1-4, (Translation accuracy, Linguistic correctness, Signing naturalness, and Recording quality), the bar chart shows the average and standard error of expert evaluation. For Q5 (Signing captured), the bar chart shows the percent of recordings evaluated as having captured the full signing space.

a page of text. It is also possible that the impact of the interface on the translation quality may vary depending on the experience or fluency of the user.

7.3 Limitations and Future Work

There is a need for a larger, more longitudinal study to see how users use the site over a period of time in the real world, rather than a short 1 hour session where they use the site for the first time and answer survey and interview questions with a researcher. Additionally, most of our participants already had a Bachelor’s degree, which may have biased our results; as a result, it is important for future studies to capture more diverse participants from the DHH community. Such studies would allow for deeper insight in user participation and behavior, and the additional data collection would enable deeper linguistic analysis and open up several research questions.

Since some participants in our user study skipped sentences, selecting nonconsecutive sentences to contribute, there are gaps in the articles. Our user study participants supported the idea of the website, said it was easy to use, but many of them said they would personally not contribute themselves. To encourage users to contribute in completeness, further research is needed to investigate different incentivization methods. There are several ways we can imagine this happening, such as strengthening the gamification inside the website (emphasizing the experience points they earn as they contribute, displaying a leaderboard of the top contributors), or monetary compensation for some arbitrary milestone of amount of ASL videos an user contributes. Another possible avenue to

investigate is educational tasks, e.g. ASL interpreting students could contribute to gain credits for certification or program requirements.

For this user study, we chose to implement a stand-alone website pre-populated with a sample of Wikipedia articles, limiting the type of content available for participants. We chose this implementation, rather than a web plug-in or other setup with broader content for several reasons: ability to choose English texts that are open for public use, utility to users in having a complete translation as opposed to sparse translations across more content, and implementation feasibility. Still, user study participants brought up many different types of content and explained their experiences with other real-world content. Consequently, other types of content, and expansive interface designs including web plug-ins, should be explored. The utility of our interface with other signed/written language pairs, or exploring other potential user groups (e.g. those recommended by participants, such as K-12 students) could also be investigated. Different use cases may or may not require other interface changes, which would be explored in this research avenue.

There were two major concerns brought up during our user study. Users were concerned about the level of control over data quality – since this is a crowdsourced approach, it is the contributors’ responsibility to have a good background in their signed videos, ensure there is good lighting, and that the video is not choppy or blurry. The other major concern was privacy. This is a very complicated topic ([25], [9], [7]), and more research is needed about privacy concerns when it comes to crowdsourced ASL datasets. Another data quality research question is whether the crowd would be able to control the data quality at a bigger scale. We included an up-vote/downvote button where participants could give feedback, but

we did not study this further, since 12 out of 19 participants did not use it. We also had a small number of sentences from each participant, but if a larger and more longitudinal study was conducted, it could be investigated how users use this feature.

We have also run a small experiment comparing the quality of translation recordings made through our interface and through a state-of-the-art setup. This exploratory study suggested that the quality of translations created through ASL Wiki are comparable to those created through state-of-the-art setups, and potentially might enable slight improvements. While this is promising, we have not evaluated the crowd-generated dataset from our participants (as we did not have a control dataset to compare to, since general community members do not normally engage in translation). It would be useful for future researchers to investigate this, as well as to conduct in-depth linguistic analysis. For instance, it is possible that our interface reduced the cognitive load of the signer, as well as the technical requirements, which may have elicited more natural and linguistically correct translations.

As mentioned above, privacy is another issue that may impact the design and use of ASL Wiki and future work. The research community has only recently begun to explore privacy concerns related to sign language videos and thought about how they can be addressed [9, 26]. This initial work began to explore the impact of filtering videos, for example by blurring the video or anonymizing facial features. However, acceptability of these approaches is poorly understood, and their technical implementation is limited. Indeed, it is possible that the community might prefer different approaches altogether, for example protective licensing or enhanced security and transparency of data use. Once a better understanding of the privacy needs and appropriate solutions have been developed, such techniques could be incorporated into ASL Wiki and similar applications, and make a ripe area for future work.

8 CONCLUSION

The lack of sign language bilingual resources and the lack of sign language datasets are difficult problems to solve, mainly due to the cost, resources needed, and amount of human effort required to label and annotate data. In this work we have addressed both of these problems by presenting a novel interface. Our interface provides a side-by-side ASL and English synchronized interface, streamlines pre-labeled data collection, and enables a crowd to contribute to piecemeal translation. We pioneer exploration of the question of how to enable everyday signers to contribute to continuous content translation efforts, and how DHH users would respond to crowd-generated content.

ACKNOWLEDGMENTS

We thank Chinmay Singh for initial work on the prototype; Vanessa Milan for design guidance; Bill Thies and Naomi Caselli for thoughtful discussions; and Philip Rosenfield, Paul Oka, and Mary Bellard for general support for the project.

REFERENCES

- [1] Oliver Alonzo, Matthew Seit, Abraham Glasser, and Matt Huenerfauth. 2020. *Automatic Text Simplification Tools for Deaf and Hard of Hearing Adults: Benefits of Lexical Simplification and Providing Users with Autonomy*. Association for

- Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376563>
- [2] Oliver Alonzo, Jessica Trussell, Becca Dingman, and Matt Huenerfauth. 2021. *Comparison of Methods for Evaluating Complexity of Simplified Texts among Deaf and Hard-of-Hearing Adults at Different Literacy Levels*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3411764.3445038>
- [3] Larwan Berke, Sushant Kafle, and Matt Huenerfauth. 2018. Methods for Evaluation of Imperfect Captioning Tools by Deaf or Hard-of-Hearing Users at Different Reading Literacy Levels. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3173574.3173665>
- [4] Jeffrey P Bigham, Chandrika Jayant, Hanjie Ji, Greg Little, Andrew Miller, Robert C Miller, Robin Miller, Aubrey Tatarowicz, Brandyn White, Samuel White, et al. 2010. Vizwiz: nearly real-time answers to visual questions. In *Proceedings of the 23rd annual ACM symposium on User interface software and technology*. 333–342.
- [5] Erin Brady, Meredith Ringel Morris, and Jeffrey P Bigham. 2015. Gauging receptiveness to social microvolunteering. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1055–1064.
- [6] Danielle Bragg, Naomi Caselli, John W. Gallagher, Miriam Goldberg, Courtney J. Oka, and William Thies. 2021. *ASL Sea Battle: Gamifying Sign Language Data Collection*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3411764.3445416>
- [7] Danielle Bragg, Naomi Caselli, Julie A. Hochgesang, Matt Huenerfauth, Leah Katz-Hernandez, Oscar Koller, Raja Kushalnagar, Christian Vogler, and Richard E. Ladner. 2021. The FATE Landscape of Sign Language AI Datasets: An Interdisciplinary Perspective. *ACM Trans. Access. Comput.* 14, 2, Article 7 (jul 2021), 45 pages. <https://doi.org/10.1145/3436996>
- [8] Danielle Bragg, Oscar Koller, Mary Bellard, Larwan Berke, Patrick Boudreaux, Annelies Braffort, Naomi Caselli, Matt Huenerfauth, Hernisa Kacorri, Tessa Verhoef, Christian Vogler, and Meredith Ringel Morris. 2019. Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective. In *The 21st International ACM SIGACCESS Conference on Computers and Accessibility* (Pittsburgh, PA, USA) (ASSETS '19). Association for Computing Machinery, New York, NY, USA, 16–31. <https://doi.org/10.1145/3308561.3353774>
- [9] Danielle Bragg, Oscar Koller, Naomi Caselli, and William Thies. 2020. Exploring Collection of Sign Language Datasets: Privacy, Participation, and Model Performance. In *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility* (Virtual Event, Greece) (ASSETS '20). Association for Computing Machinery, New York, NY, USA, Article 33, 14 pages. <https://doi.org/10.1145/3373625.3417024>
- [10] Danielle Bragg, Kyle Rector, and Richard E Ladner. 2015. A user-powered American Sign Language dictionary. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1837–1848.
- [11] B Cartwright. 2017. Signing Savvy. <https://www.signingsavvy.com/>
- [12] ASL Clear. 2022. <https://aslclear.org/>
- [13] Helen Cooper, Eng-Jon Ong, Nicolas Pugeault, and Richard Bowden. 2012. Sign Language Recognition Using Sub-Units. *J. Mach. Learn. Res.* 13, 1 (jul 2012), 2205–2231.
- [14] ASL Core. 2022. <https://aslcore.org/>
- [15] Amanda Duarte, Shruti Palaskar, Lucas Ventura, Deepti Ghadiyaram, Kenneth DeHaan, Florian Metzke, Jordii Torres, and Xavier Giro-i Nieto. 2021. How2Sign: A Large-scale Multimodal Dataset for Continuous American Sign Language. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [16] Karen Emmorey, Chuchu Li, Jennifer Petrich, and Tamar H. Gollan. 2020. Turning languages on and off: Switching into and out of code-blends reveals the nature of bilingual language control. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 46, 3 (2020), 443–454. <https://doi.org/10.1037/xlm0000734>
- [17] Jens Forster, Christoph Schmidt, Thomas Hoyoux, Oscar Koller, Uwe Zelle, Justus Piater, and Hermann Ney. 2012. Rwth-phoenix-weather: A large vocabulary sign language recognition and translation corpus. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. 3785–3789.
- [18] Ann Grafstein. 2002. HandSpeak: A Sign Language Dictionary Online. <https://www.handspeak.com/>
- [19] Julie Hochgesang, Onno Crasborn, and Diane Lillo-Martin. 2022. Sign Bank. <https://aslsignbank.haskins.yale.edu/>
- [20] Leala Holcomb and Jonathan McMillan. 2022. <http://www.handsland.com/>
- [21] Matt Huenerfauth and Vicki Hanson. 2009. Sign language in the interface: access for deaf signers. *Universal Access Handbook*. NJ: Erlbaum 38 (2009), 14.
- [22] Deaf Studies Digital Journal. 2009. <https://www.deafstudiesdigitaljournal.org/>
- [23] Walter Lasecki, Christopher Miller, Adam Sadilek, Andrew Abumoussa, Donato Borrello, Raja Kushalnagar, and Jeffrey Bigham. 2012. Real-time captioning by groups of non-experts. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*. 23–34.

- [24] Walter S Lasecki, Christopher D Miller, Raja Kushalnagar, and Jeffrey P Bigham. 2013. Legion scribe: real-time captioning by the non-experts. In *Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility*. 1–2.
- [25] Sooyeon Lee, Abraham Glasser, Becca Dingman, Zhaoyang Xia, Dimitris Metaxas, Carol Neidle, and Matt Huenerfauth. 2021. *American Sign Language Video Anonymization to Support Online Participation of Deaf and Hard of Hearing Users*. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3441852.3471200>
- [26] Sooyeon Lee, Abraham Glasser, Becca Dingman, Zhaoyang Xia, Dimitris Metaxas, Carol Neidle, and Matt Huenerfauth. 2021. American Sign Language Video Anonymization to Support Online Participation of Deaf and Hard of Hearing Users. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*.
- [27] Colin Lualdi. 2022. Sign School. <https://www.signschool.com/>
- [28] Matt Malzkuhn and Melissa Malzkuhn. 2022. The ASL App. <https://theaslapp.com/>
- [29] The Daily Moth. 2022. The Daily Moth. <https://www.dailymoth.com/>
- [30] National Association of the Deaf. 2022. Position Statement On ASL and English Bilingual Education. <https://www.nad.org/about-us/position-statements/position-statement-on-asl-and-english-bilingual-education/>
- [31] Eng-Jon Ong, Helen Cooper, Nicolas Pugeault, and R. Bowden. 2012. Sign Language Recognition using Sequential Pattern Trees. *2012 IEEE Conference on Computer Vision and Pattern Recognition* (2012), 2200–2207.
- [32] Eng-Jon Ong, Oscar Koller, Nicolas Pugeault, and Richard Bowden. 2014. Sign Spotting Using Hierarchical Sequential Patterns with Temporal Intervals. In *2014 IEEE Conference on Computer Vision and Pattern Recognition*. 1931–1938. <https://doi.org/10.1109/CVPR.2014.248>
- [33] Jingnan Peng. 2020. Bringing light to the news, for those who can't hear it (video). <https://www.csmonitor.com/The-Culture/2020/0731/Bringing-light-to-the-news-for-those-who-can-t-hear-it-video>
- [34] Manaswi Saha, Michael Saugstad, Hanuma Teja Maddali, Aileen Zeng, Ryan Holland, Steven Bower, Aditya Dash, Sage Chen, Anthony Li, Kotaro Hara, et al. 2019. Project sidewalk: A web-based crowdsourcing tool for collecting sidewalk accessibility data at scale. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [35] Elliot Salisbury, Ece Kamar, and Meredith Morris. 2017. Toward scalable social alt text: Conversational crowdsourcing as a tool for refining vision-to-language technology for the blind. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, Vol. 5. 147–156.
- [36] Elliot Salisbury, Ece Kamar, and Meredith Ringel Morris. 2018. Evaluating and Complementing Vision-to-Language Technology for People who are Blind with Conversational Crowdsourcing. In *IJCAI*. 5349–5353.
- [37] Zed Sevcikova Sehyr, Naomi Caselli, Ariel M Cohen-Goldberg, and Karen Emmorey. 2021. The ASL-LEX 2.0 Project: A Database of Lexical and Phonological Properties for 2,723 Signs in American Sign Language. *The Journal of Deaf Studies and Deaf Education* 26, 2 (02 2021), 263–277. <https://doi.org/10.1093/deafed/ena038> arXiv:<https://academic.oup.com/jdsde/article-pdf/26/2/263/36643382/ena038.pdf>
- [38] Ather Sharif, Paari Gopal, Michael Saugstad, Shiven Bhatt, Raymond Fok, Galen Weld, Kavi Asher Mankoff Dey, and Jon E. Froehlich. 2021. Experimental Crowd+AI Approaches to Track Accessibility Features in Sidewalk Intersections Over Time. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–5.
- [39] Carol Bloomquist Traxler. 2000. The Stanford Achievement Test, 9th Edition: National Norming and Performance Standards for Deaf and Hard-of-Hearing Students. *The Journal of Deaf Studies and Deaf Education* 5, 4 (09 2000), 337–348. <https://doi.org/10.1093/deafed/5.4.337> arXiv:<https://academic.oup.com/jdsde/article-pdf/5/4/337/9835826/337.pdf>
- [40] Wikipedia. 2001. <https://www.wikipedia.org/>
- [41] Wikipedia. 2022. Deaf News. https://en.wikipedia.org/wiki/Deaf_News
- [42] Alicia Wooten and Barbara Spiecker. 2022. <https://www.atomichands.com/>

A APPENDIX

A.1 Semi-structured user study interview questions

Below is the semi-structured interview questions that were discussed with participants as part of the user study:

- Role/relation to ASL: What's your role/relationship with ASL? (e.g. native speaker, primary language, ASL teacher, use ASL at work, etc...)

Reading

- Did you primarily look at the ASL or English part? [Follow up to estimate percentage (0% ASL 100% English vs 100% ASL 0% English)]
- How did viewing different signers affect your experience? (If applicable)
- On a scale from 1-5 (1- very difficult), how understandable was the ASL content you viewed? Can you explain why you chose this number?
- On a scale from 1-5 (1-very difficult), how understandable was the English content you viewed? Can you explain why you chose this number?
- Was it helpful to view the content in both English and ASL? Why or why not?
- Did you use the upvote/downvote feature? Why or why not?
- How easy was the interface to use? (1-5: Very difficult – Very easy) If difficult, did information overload contribute to difficulties?
- What did you like or dislike about the interface?

Recording

- Did you find any content challenging to record? If so, what made it challenging?
- Did you use any strategies while recording content? If so, what were they?
- On a scale from 1-5, how easy was the interface to use? (1-5: Very difficult– Very easy)
- What did you like or dislike about the interface?

Desirability

- Do you wish more content online provided both English and ASL? (1-5: Strongly disagree– Strongly agree) If so, can you give some examples of when you wanted content provided in both languages?
- Would you be interested in generating content available in both English and ASL? (1-5: Strongly disagree– Strongly agree)
- What benefits do you feel this site offer to you as a user, if any?
- What concerns do you have in using a website like this, if any?
- How enjoyable was using the website, overall? What did you like/dislike?
- Would you want to use a website like this to read content in the future? Why or why not? Is there different content you would want to read (e.g. movie scripts, podcast, etc.)?
- What type of *Wikipedia* content would you want translated (i.e., picking from the list of topics on the Wikipedia landing page – food, math, Deaf culture, etc.)?
- Would you want to use a website like this to contribute recordings in the future? Why or why not?
- How likely are you to recommend this website to others? (1-5: Very likely – Very unlikely) If so, who would you recommend this to, and for what purpose (e.g. ASL students for learning, people with certain English/ASL fluency, etc.)?