A Counting for Silence

by

Nina M Lutz
B.Sc Computer Science with Design, MIT

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of

Master of Science
Master of Science in Media Arts and Sciences

at the
Massachusetts Institute of Technology
September 2021

Author
……..
Nina M Lutz
Program in Media Arts and Sciences, 8-20-21

Certified by
……...……
Ekene Ijeoma
Assistant Professor, Media Arts and Sciences

Accepted by
……
Tod Machover
Academic Head, Program in Media Arts and Sciences

A Counting for Silence

by

Nina M Lutz

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, on 8-20-21 in partial fulfillment of the requirements for the degree of

Master of Science
Master of Science in Media Arts and Sciences

Abstract:

A Counting is a project that asks the question of who gets counted and how. Created in response to the undercount of marginalized communities in the US Census, *A Counting* was released in 2020 as a participatory, generative voice portrait providing representation of the linguistic diversity of the United States. *A Counting: Sign Language* is an edition of this work that furthers this inquiry with regard to the role of participatory art in representation, in particular the representation of sign language.

My thesis project encompasses the technical contributions of a video-first recording interface, a transcription interface, and a computer vision process to segment user recordings into discrete numbers and into a final portrait of 1-100. Additionally, this thesis encompasses theoretical groundings and histories of participatory public artworks that utilize technology and video, positioned with disability aesthetics and linguistic justice, particularly with regard to sign language.

Overall, this thesis hopes to continue to build on the series of works in Poetic Justice and methods for enacting such works, and bring about new vocabularies and analysis for exploring representation through participatory digital art.

Thesis advisor:
Ekene Ijeoma
Assistant Professor, Media Arts and Sciences

A Counting for Silence

by

Nina M Lutz

This thesis has been reviewed and approved by the following committee members

Ekene Ijeoma ……...……

Assistant Professor, Media Arts and Sciences
Massachusetts Institute of Technology, Media Lab

Caroline Jones, PhD…….

Associate Dean for Strategic Initiatives, School of Architecture + Planning, MIT
Professor, History Theory Criticism Section, Department of Architecture, MIT
Massachusetts Institute of Technology, Department of Architecture

Zach Lieberman …...……

Adjunct Associate Professor, Media Arts and Sciences
Massachusetts Institute of Technology, Media Lab

1. Introduction

Throughout my studies, I have been fascinated by the idea of using technology to make us think about other people -- particularly people who are different from us. This comes perhaps as a rebellion from most computer science training, where the hottest thing seems to be making computers think more like people. Perhaps the rest is a rebellion from my aesthetic training, where I long since considered the cultural criticism against the “normative” in visual communication and beauty of the human form and identity.

My particular technical field of inquiry has always been computational vision and imagery -- particularly geometric foundations and methods. Much of my original work at the start of my degree program involved interactive artworks with participants and faces and lights. These works varied vastly from my early research at the Media Lab with Ira Winder when we focused on projects of complex simulations to colloidal lighting fixtures or to makeup models that sought to enable individuals to affirm their gender identity.

Much of this thesis is touched with kismet. I went to a film screening with the conversation facilitated by Caroline Jones in 2019 whilst working on a project around cultural identity, archives, and methods for 3D dynamic projection mapping. At the time I had no idea Caroline would be later reading my thesis document, and that project would not be my thesis.

I have a pin from Ekene, given to me when I was a student in a different group, which lived in my office in that group space. Having not a clue that he would be my advisor upon graduating and that I would graduate from this group. It is now on its way to Seattle, in a UHAUL, with the rest of my things, many brought here in 2019 from Arizona as I originally thought I would be in Cambridge until 2025 and leaving as Nina Lutz, Ph.D. The “Nina Lutz, Ph.D.” project is on hold right now, but I know it will come back online at some point in the near future.

This project itself exists because of an inquiry from a user regarding sign language support of the preceding voice edition of this artwork, which I was a developer on.

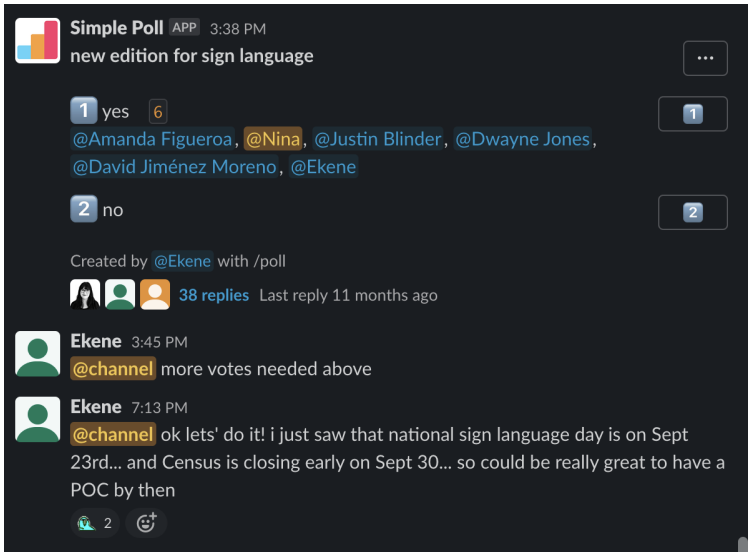


Figure: This entire thesis started with a Slack vote seen above that was prompted by an email Ekene received.

Just as I was struggling to find a new thesis idea to pursue fully digitally after a particularly difficult first year in this program of crises and a pandemic, and after 5 years of working with both physical and digital projects. It felt like a natural fit, given my interest and previous work in computer vision and passion for visual media enabled by computation.

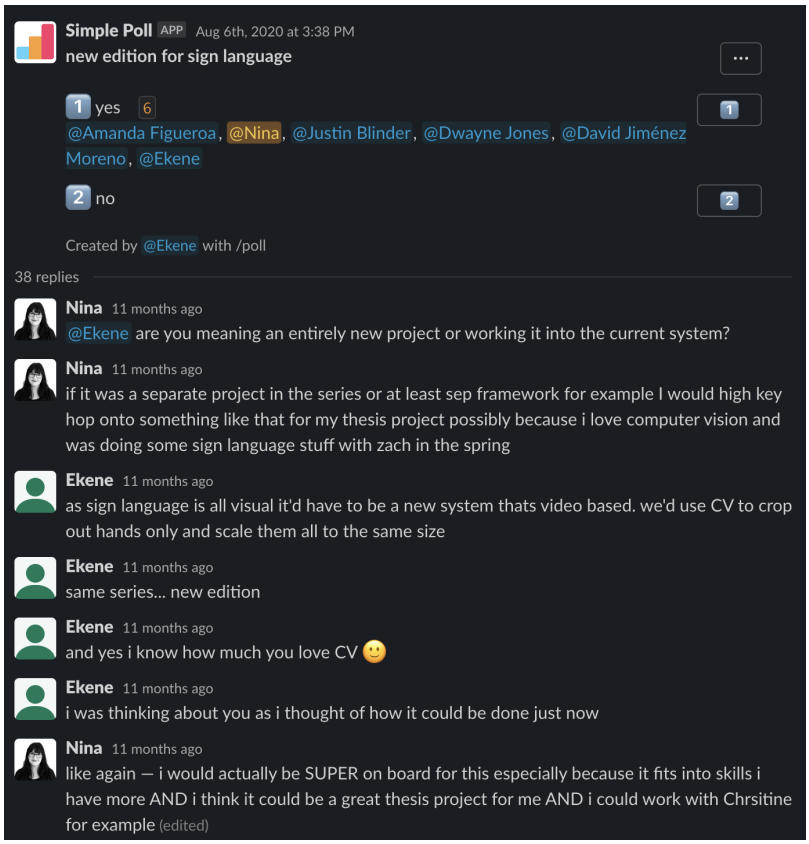


Figure: Additionally preliminary discussions of this thesis on Slack

While the project was forming, it was another bit of luck perhaps, that Zach Lieberman had a close friend from his schooling who is a Professor at Gallaudet University, and was kind enough to interface with me through this work.

Throughout this document, I have tried to incorporate my learnings and a sense of the vast interconnections that works and histories it engages with -- partially out of my own academic ambitions, but partially as a testament to the idea of a thesis produced so saliently from so many small circumstances that were perhaps breadcrumbed all the way.

The thesis first introduces the *A Counting* series and motivation for this project *A Counting: Sign Language*, then the histories and art forms the work engages with. The next two sections outline my contributions, both quantitative and qualitative. Finally, a section engages with the evaluation and ethics of the contributed work.

This wasn't the thesis I thought would occur, for this degree. But I have learned much from its research, development, and now documentation process. And I hope it will be but one of the works I will do throughout my career that seeks to produce new interfaces and lines of inquiry that push and open visual communication and artistic projects to be more inclusive and equitable in the future.

2. Existing Work and Motivation

A Counting was released in 2020 as a generative, crowdsourced voice portrait of participants counting from 1-100 in their languages with a different language and voice for each number. Currently, there are over 700 calls in over 90 languages. Participants call a hotline and count 1 to 100 in a language of their choice. These audio recordings are then computationally sliced with signal processing into distinct number files, which are then transcribed by participants before being added to the artwork.

As on the website at the beginning of this work:

> A Counting is an ongoing series of video and sound-based portraits of the US. Created in response to the US Census which has historically misrepresented the ethnic and linguistic diversity of the US, A Counting serves as a meditation and speculation on what a truly united society would sound and look like.
> [a-counting.us, 2020 October]

However, there are thousands of undercounted individuals who communicate with sign language, which is also underrepresented in censuses [Mitchell], as made evident by movements through history. In a world where communication is power, this thesis centers itself on holding space for those who do not communicate with spoken languages.

This thesis encompasses a signed edition of the *A Counting* artwork -- an online, participatory portrait of individuals signing 1 to 100 in any sign language, with a different hand for each number. This includes the recording and transcription interfaces for these signed numbers, a method to process the individual signed numbers, and the final video portrait. Now, a-counting.us says the following after this thesis:

> A Counting is an ongoing series of linguistic portraits of the US featuring various spoken and sign languages. Created in response to the US Census which has historically misrepresented the ethnic and linguistic diversity of the US, A Counting serves as a meditation and speculation on what a truly united society would look and sound like.
> [a-counting.us, 2021 July]

While the artifact of this project is a video portrait and codebase, this thesis will also present an analysis of relevant histories and precedents, as well as qualitative contributions of new frameworks and theories and educational contributions.

This thesis continues to build on the series of participatory works in the Poetic Justice research group and brings about new vocabularies for analytical and pedological framing regarding representation and participation in technologically enabled artworks.

Having *A Counting* support sign language incorporates another powerful statement against the systemic miscount of people in the United States and shows unity across the power of forms of communication in a more just count.

It is the hope that projects and analyses like this one will inspire more technologists and artists to be more critical in the justices, media, and identities that they engage with.

3. History and Context

It is worth noting that *A Counting: Sign Language* is a public online participatory artwork. In particular, it is a video linguistic portrait. This brings about a rich historical and artistic content to this work: participatory artwork, conceptual artwork, public artwork -- as well as traditions in web, language, and video artworks.

Furthermore, this work draws from and engages with disability aesthetics, cultures of signed languages and Deaf communities, and linguistic and societal justice.

This section will try to provide an overview and semi-rigorous analysis of these relevant academic spheres and artistic and social practices.

3.1 Conceptual art

> ***"The 'purest' definition of conceptual art would be that it is an inquiry into the foundations of the concept 'art'" - Joseph Kosuth***

A Counting is a series of conceptual artworks. Within the school of conceptual art, concepts take precedence over materials and aesthetic artifacts [MoMA]. This is not to say that the aesthetics and mediums of the multimedia portrait of *A Counting* are not important. The vocal edition is very carefully designed to have even timing between audio files and consistent spacing of the text. The signed edition went through many iterations to ensure compositional consistency and good design practices from aspect ratios to color balancing, outlined in this thesis extensively.

But it is the concepts that ground the framing, execution, and experience that participants have in *A Counting.*

3.1.1: History and methods

> ***Almost every work of serious contemporary art recapitulates, on some explicit or implicit level, the historical sequence of objects to which it belongs.***
> ***...***
> ***Consciousness of precedent has become very nearly the condition and definition of major artistic ambition.***
>
> ***[Thomas Crow, 1995, from Buchmann p 53]***

Conceptual art rose to popularity in the mid 20th century, emerging formally with Henry Flynt coining the term in 1961 [Almenberg, 2010], but becoming more accepted and popular in galleries and the art world in the mid-1960s [Almenberg, 2010]. Despite these origin timelines, Marcel Duchamp is widely regarded as the predecessor of conceptual art with his readymade artwork Fountain in 1917 [Buchmann, 2006].

Soon conceptual art started to be critiqued and defined. One such group within the scholarship of conceptual art was Art & Language, a conceptual artists collaboration that wrote several publications on the matter [Buchmann, 2006].

> Charles Harrison, the editor of the journal *Art-Language,* laid down the requirement for any Conceptual art aspiring to critical interest that it conceive a changed sense of the public alongside its transformation of practice.
>
> ….
>
> If the history of Conceptual art is to maintain a critical value in relation to the apparent triumph of visuality, it must meet the condition implied in their judgment on its fate
>
> 1. It must be living and available rather than conclude;
> 2. It must presuppose, at least in its imaginative reach, renewed contact with lay audience;
> 3. It must document a capacity for significant reference to the world beyond the most proximate institutions of artistic display and consumption.
>
> [Thomas Crow, 1995, from Buchmann p 56]

By some definitions, artwork has always been conceptual. Many artists throughout human history sought to convey a concept or provoke rigorous thought in the viewer. The key distinction is that conceptual art is that the whole goal, purpose, and critique of the art lies in its concept, departing from the traditional ideas of art as an object first and foremost:

> In conceptual art the idea or concept is the most important aspect of the work. When an artist uses a conceptual form of art, it means that all of the planning and decisions are made beforehand and the execution is a perfunctory affair.
> [LeWitt, ‘Paragraphs on Conceptual Art’, Artforum Vol.5, No.10, Summer 1967, p 79-83]

This history is important to consider regarding the social movements at the time that were also occurring, such as the American Civil Rights movement and, regarding this project, the American Disabilities movement, including Deaf culture’s start and the starting of Deaf culture schools of thought and art [Hendron, 2020, p 106]. Conceptual art was a push against the highly commercialized art forms at the time, such as 1950s pop art, as well as a response to the changing society and civil rights movements of the 1960s [Buchmann, 2006].

Conceptual art, by its definition, has many methodologies and common motifs. Relevant to this thesis, language is one such medium and method and has long been a significant tool in conceptual art [MoMa]. While there are many directions and views of language within conceptual art, three main categories stand out: language as delivery of artwork, imagery and symbolism, and juxtaposition of language and imagery and objects.

In her essay “Language, Reality, Irony: The Art Books of Jaroslaw Kozlowski”, Luiza Nader explores how conceptual artist Jaroslaw Kozlowski uses language and the relationship of language and truth and grammar as structure and rules towards his artistic works, particularly his art books [Buchmann]:

> [Kozloski]...treats language as an autonomous base from which one does not look for the truth, but instead, for freedom. His reflections on language, on its relation to the world, and on the condition of the subject imprisoned within the language and language games, force us to regard Conceptualism from a subjective perspective, with doubt and distance. [Luiza Nader, 2006 from Buchmann, p 102]

This treatment and interpretation of language suggest language not just as representation but also as something that enables life while having a life of its own. This brings about ideas of agency around language and its influence. Kozlowski pushed this intrigue of the power of language through explorations of numbers and grammar, in particular, to introduce paradoxes within these more systemic aspects of language [Buchmann, 2006, p 103]. This is apparent in *Deka-log*, which “depicts multiplication of signs according to their numerical value; a single one being followed by two twos and three threes, etc.” [Buchmann, 2006, p 103].

These books were a perfect artifact for many of Kozlowski’s musings around language, many of which were born of philosophy as opposed to artistic precedents, making language a perfect medium as it “...was not seen as either abstract or realistic. The artist simply suspended any and all categorizations.” [Buchmann, 2006, p 116]. This brings back to the central ideas of language as freedom, a fitting concept to convey with language itself as a medium.

Along with language as the artifact as a medium, participatory drawings, delivered by language in instructions, is another significant contribution towards conceptual art practice. Most prolific, perhaps, are many of the drawings of Sol Le Witt. Many of Le Witt’s drawings challenged the concepts of artistic creation and authorship [Buchmann, 2006]. These drawings were delivered in instructions that resulted in the work (and often resulted in different interpretations and thus different drawings) [Bloom, 2019]. Le Witt utilized language to give his drawings new conceptual groundings in challenging what it means to create a drawing but also to create new works through relying on the interpretation of these instructions by others.

Joseph Kosuth’s “In One and Three Chairs” is an installation of a physical chair, an image of it, and the dictionary definition of the chair -- raising questions of different states and truths of the chair [MoMa]. The chair itself exists in the real world, but language and its definition give it meaning in society. Kosuth once said, “art is making meaning” [Joseph Kosuth, quoted in Christopher Lyon, ed., Contemporary Art in Context (New York: The Museum of Modern Art, 1988), p 47.].

Aristi Barbara Kruger expands upon the idea of words and art-making meaning -- in exploring that through this meaning, words have power over us:

> I work with pictures and words because they have the ability to determine who we are and who we aren't. [Kruger, Artist Website]

Kruger uses written language, typographic styling, and imagery with a sense of urgency in collage and with the incorporation of cultural points and phrases directed at and involving the viewer as a spectator:

> Kruger addresses media and politics in their native tongue: sensational, authoritative, and direct. Personal pronouns like "you" and "I" are staples of Kruger's practice, bringing the viewer into each piece. "Direct address has motored my work from the very beginning," [Barbara Kruger - Bio, The Board]

Kruger also uses culturally relevant declarative phrases in her artwork to encapsulate another power of language -- the power of direct addressing and bringing someone into a dialogue or, in Kruger's case, accusing someone.

The diverse methods by which language can function as a medium and delivery tool of conceptual art speak to the many facets of language and its concepts that can be pulled into artwork.

3.1.2: Concepts in *A Counting*

These multimedia portraits convey a unified count of linguistic diversity in the United States and create space to experience and reflect on this count. The key inquiry being: Who counts and why? From a conceptual framework, there are many layers of ideas, abstractions, and provocations that the *A Counting* series seeks to engage with through this line of inquiry. Below is a summary of some of the salient concepts within *A Counting*. Language and counting are elaborated on in the next section:

Table: Summary of concepts in the *A Counting* series

Concept	**Description**	**Artistic embodiment**
Acknowledgment	Acknowledgment is the acceptance of a truth, and often the first part of recovering from trauma or grief. Additionally, to acknowledge someone is to validate their reality and struggles.	All editions hope to have a “language acknowledgment” of an indigenous language as the first number of the generative count. Holding space for rarer languages with silences or blanks that do not possess recordings of all of their numbers, done with silence over transcripts for the Lenape tribe in the voice edition. Acknowledging the first iteration of this work did not accommodate signed languages, and releasing this very edition.
Meditation	The act of counting has been described as “soothing” by some participants. It is a rhythmic task that people do not often have to perform.	The interface of both signed and spoken editions are designed for participants to be able to start and stop counting when they are ready so that they can be present and relaxed in their experience of counting.
Unity	In English, “unity” is the state of being united or a whole.	The idea of combining multiple languages and people in one portrait -- where people of all different backgrounds and locations are part of the same count. Participants in this model are also united in their efforts to accomplish a whole count of multiple languages.
Equity	Equity implies an understanding of just and fair treatment, in the context of what that group may or may not need.	Rarer languages in the system are inversely weighed during count generation so that they will appear more often.

3.1.3: Counting and Language

Who counts and why? This is a central question within the series that *A Counting: Sign Language* takes a step further in acknowledging that in the first edition, signed languages were not counted.

The portrait itself, the output of the artwork, is a generative, infinite sequence of counts of 1 to 100. It’s in the name -- the project is a **counting**, playing on the word accounting as it accounts for the miscounts of marginalized people.

Counting as a site for quantization of human populations. Counting as a representation of language. Counting as a unifying methodology to stitch together multiple participants. Counting as an allegory for percentages and statistics. These and more are lenses that could be meditated on through this artwork.

No one can count forever. Mathematically, no one could even count out the infinite numbers between 1 and 2 (1.00000000000001, 1.00000000000000011, 1.00002, 1.3, etc, etc). These counting portraits are generative in part because every count is unique and the sequence keeps going, with new portraits generated each hour. The count is constantly changing and, like real numbers, becomes infinite as more and more people contribute.

Not everyone's voice is guaranteed to be counted in many systems, but in *A Counting* everyone gets the chance to count and be counted as a contributor to the artwork. This concept of counting as attribution and acknowledgment is an additional important conceptual layer to the framing of the participation and presentation of this project.

Perhaps most importantly, the choice of counts as a site of language representation is another concept that this portrait interrogates. Counting is one of the first things someone learns during language acquisition.

This is especially true for individuals who do not speak a language fluently. This opens the portrait to more individuals who may not know the entire language that they wish to represent, often due to societal pressures, past eradication of the language, and other personal and familial reasons, such as access. This is a major point within sign language, as many individuals do not learn to sign.

This is part of the motivation for how this edition is framed -- this edition is meant for users of any signed language, not just ASL, and not just Deaf users. However, this thesis acknowledges that it is impossible to have a conversation about signed languages in the United States without discussing Deaf culture and individuals who are hard of hearing or deaf.

Within the Deaf community, in particular, many members have faced language deprivation. Many are born to hearing families that do not learn or teach their children sign language -- some may not have access to sign language resources and others may believe that signed languages are inferior and therefore push for their children to communicate in other ways.

This can lead to the phenomenon of language deprivation, where a child is not exposed to linguistic concepts early enough in their cognitive development, affecting them later in life [Cheng, et. al, 2019].

> Under certain conditions, neither spoken language nor sign language is accessible to deaf children resulting in early language deprivation. Thus, congenitally deaf individuals vary in their early language environment, offering a rare opportunity to investigate the role of early language experience in the development of the neural language network.

> Research on this question is also crucial to raise the awareness of the potential negative sequelae of early language deprivation in deaf children. [Cheng, et. al., 2019]

Deaf individuals not having access to sign language is due to a culmination of historical discrimination against singed languages and systematic efforts to eradicate them, such as oralism [Hendron, 2020, p 107]. Other linguistic minorities, such as indigenous peoples, may also face the consequences of language eradication.

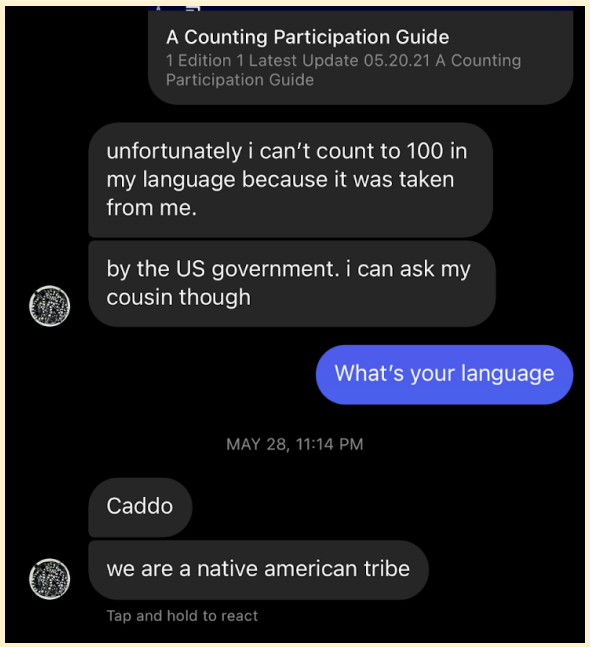


Figure: Screen capture of an individual's response to a Call to Action for *A Counting* of an individual who does not know their ancestral tongue due to discriminatory and eradicationist practices against indigenous languages.

At the intersection of these are indigenous signed language systems, which are exceedingly rare due to genocide, colonization, and displacement of indigenous peoples, including oralism and residential schooling of indigenous children [Davis, 2017] [Hilleary, 2017].

This shows the importance of having not only the signed edition of this artwork but also having the language acknowledgment at the start of each count, repeating infinitely, for both the spoken and signed editions.

3.1 4: Strengths, critiques, and challenges of conceptual art

> ***"Modern art history and art museums created a highly controlled system that divides our visual heritage into two binary categories: art, and everything else." - [Lev Manovich, 2020, p 98]***

Since these artworks surround ideas as the center of their treatise of work, there are often issues regarding the market value of such works [Almenberg, 2010]. This devaluation is not only restricted to conceptual art but artifacts that have much to do with heritage and personal

significance (or, concepts) but are not deemed valuable by the art world at large [Manovich, 2020, p 98]. This creates a marketing issue for conceptual artists, and may explain why it was not as popular during previous eras where art objects were being sold for high value [Almenberg, 2010].

It is perhaps worth noting that conceptual and participatory artwork share this marketability issue and it is through doing artworks like these in institutions and collectives that they can be made more possible.

While there is space for criticism regarding the fact that in some ways conceptual art can be easier to enact -- it is easier to place a chair, a poster, and a photograph of the chair in a room than to paint a mural -- much of conceptual art's talent comes from the writing, research, framing, and presentation of the work. This stands to reason in that, by many historical definitions, conceptual artwork should be analyzed on the grounds of the inquiry it provokes -- not just the artifact and artistic sense of said artifact.

Conceptual art is uniquely suited to raise critical dialogue in the arts and society at large about important issues and aspects of the human experience. Much of this thesis examines the relationship between aesthetics, history, and culture and how these may enact change upon one another.

And while conceptual art and related art forms can raise important issues through art, there are critiques regarding the sincerity of raising these issues in the setting of elite institutions and if this contradicts the artistic contributions of the concept itself.

> "Rather, conceptual artworks and those that derive from them provide an understanding (gained only through close attention to the specificity of those works) of the manner by which culture becomes stratified, and hence offer privileged access to the potential and actuality of ambitious contemporary art." - Alexander Alberro, quoted by Bishop, 2012

While conceptual art exists within the very privileged systems that it may critique, it is important to note that conceptual art is not a monolith. All conceptual artworks have different societal critiques and different framings, but also that the exhibition of conceptual artworks, even those that are critical of elite institutions, may add to its strengths in many framings. *A Counting* is one such artwork by placing counts of anyone in gallery settings while also keeping this same artwork free and public online.

The irony of many of these weaknesses is that many arguments against conceptual art are grounded in the difficulty of quantifying conceptual artworks and questioning their motivations -- the very key concept that this artwork itself is built upon by considering who counts and why.

3.2 Participatory artwork

> ***"Participatory Art is founded on a vision of human creativity being intrinsically of great interest and so in itself worth pursuing for people in order to gain a deeper understanding of what it means to be a human being" - Gustaf Almenberg***

The entire *A Counting* series is by nature a participatory artwork, meaning that art itself is incomplete, or in this case nonexistent, without participation. In much literature, the word "spectator" is used to describe this relationship between an individual that is not the artist contributing to the work [Almenberg, 2010] [Bishop, Artificial Hells, 2012].

Much of this comes from the idea that the common denominator of participation in the artwork is viewing and interpreting the artwork through one's mind, described as the "spectator's imagination in the work of art" [Almenberg, 2010, p 15].

This interpretation and participation may solicit ideas of "interactive" artwork. Particularly for digital works such as *A Counting.* This grammar comes from the nature of multimedia websites and the importance of interaction design within online experiences.

While participants do need to interact with the recording and transcription interfaces of *A Counting* to participate, this is not an interactive artwork.

Instead, the unique role and action potential of participation in this work is beyond interaction -- participants are actively creating artwork through an experience that is enabled by web technologies. This distinction means that this is not an interactive web artwork but an online participatory work. Additionally, this project is edited and curated in its entirety, similar to how participation in anthropology and documentary practices is handled.

Using frameworks from critics and practitioners, this section will provide historical and artistic precedents, including salient distinguishing features between participation and interaction in the artwork. As well as outlining the feedback that participatory frameworks have with the conceptual and creative process of the artwork itself. Finally, this section will detail critiques and strengths of participatory artwork at large and the participation model of *A Counting.*

3.2.1: History and methods

Both Gustaf Almenberg and Claire Bishop, scholars who have studied and written on participatory art, place their origins sometime at the beginning of the 20th century.

Bishop accredits Italian Futurism as the first form of participatory art:

> "...first concerns Italian Futurism's break with conventional modes of spectatorship, its inauguration of performance as an artistic mode, addressing a mass audience for art, and its use of prevocational gestures (both onstage and in the streets) to increasingly overt political ends." [Artificial Hells, pg 41]

The next step of sorts was the Dadaists and the ready-mades of Marcel Duchamp, in the late 1910s and early 1920s. These *Ready-mades* were a new form of art in that Duchamp drew inquiry to the spectator's ability to "bring something to anything they look at", particularly everyday objects found by chance and their entrance into art galleries [Almenberg, 2010 p 37].

After WWI the Dada movement took off, and the emphasis of much of this movement was on personal choices [Almenberg, 2010, p 43]. While not a direct evolution, the Dadaism traditions, and characteristics influenced the paradigm and characteristics of participatory art:

> "Both kinds of art, to some extent, also made the relationship between art and spectator a central issue." [Almenberg, 2010, p 44]

Dadaists often wrote regarding

> "..the expression of "participatory theatre" to describe the kind of spectator involvement the Paris Daists elicited from their audiences...a "radical rethinking of art making" [Almenberg, 2010, p 44]

This participatory theatre component is also included in Bishop's historical analysis.

The Dadaist movement probed the interaction of art and the spectator and paved the way for serendipitous art and methods for interacting with audiences. Soon, large mobile and kinetic art became a theme of the 1920s art world, adding new forms to invoke this new relationship [Almenberg, 2010]. This resulted in a variety of art that necessitated a "physical contribution from the spectator...in turn led to choice" [Almenberg, 2010, p 70] as well as the artists of the time beginning to challenge and reconsider notions of the value of authorship and to gain an affinity for work that was not complete, but rather more open.

It was near the beginning of the Modern Movement that Marcel Duchamp and Man Ray "put into question the paramount value of the unique work of art, of the "original work of art, for the communication of the artist's ideas and intentions." [Almenberg, 2010, p 75]. This was first seen with ready-mades, but this grew to an important part of participatory and conceptual artworks in the next few decades.

In the 1950s and 1960s, the nature of openness became the next discourse leading up to the modern notion of participatory art. Umberto Eco contributed some of the more cited thoughts around this, particularly in his piece "The Open Work (1962)" and accompanying text "The Poetics of Open Work". Eco proposes an analysis and understanding that purposeful openness and incompleteness in art has intrinsic value in how it invites "mental collaboration of the consumer, who must freely interpret an artistic datum" [Bishop, 30] and additionally relates concepts of openness and possibility in the art to a reflection of scientific thought and excitement of the times [Bishop, 2006, p 32]. Eco goes as far as to claim that:

> The poetics of the ‘work in movement’ (and partly that of ‘open’ work) sets in motion a new cycle of relations between the artist and his audience, a new mechanics of aesthetic perception, a different status for the artistic product in contemporary society. - [Umberto Eco, quoted in Bishop, 2006, p 39]

This mechanic Eco writes regarding the participation of the spectator in mental interpretation ties well to conceptual art, which was also developing at this time. Almenberg argues that these parallel developments of conceptual and participatory art occurred not just because of cultural exchange and exterior social movements but also because:

> “...Conceptual Art, like Participatory Art, in a serious way set out to expand the role of the spectator...with Conceptual Art...the spectator was now asked, as an integral part of the art work, to read, to figure out, to contextualize, and to respond to statements of various kinds.” [Almenberg, 2010, p 140]

Like participatory art, conceptual practice values the conception and frameworks of the artwork and not just the final product but what space and possibilities it may enable from the spectator. This allows the two to work cohesively in multiple artworks, such as the *A Counting* series.

Despite these similarities and this compatibility, it is important to distinguish between the two. Within conceptual art, the control of the final work is not necessarily released by the artist -- this is however an integral trait of participatory art and its activation and enablement of participation in creating the artwork. Moreover, conceptual art seeks to create mental engagement while participatory art seeks to create an experience and this opportunity for creativity -- while these can coexist, they are distinct from one another.

> Conceptual Art produces the ideas to the spectator. Participatory Art stresses the involvement of the spectator through physical participation in the creative process, including even the formation of the idea or part of the idea. [Almenberg, 2010, p 145]

It stands to reason that participatory art can work to enact the ideas and frameworks that conceptual art may set out, as long as the participation is still active and consequential to this creation and additionally the concepts and conclusions presented.

In the 1990s, interactive and relational artworks prompted Nicolas Bourriaud’s 1998 essay “Relational Aesthetics” which has since become a seminal text regarding the school of thought, particularly in understanding that relational artworks center the schema that the artist intends to invoke instead of the spectator or artist:

> “It is born of the observation of the present and of a reflection on the destiny of artistic activity. Its basic hypothesis -- the sphere of human relations as site for the artwork -- is without precedent in the history of art, even though it can of course be seen” [Nicholas Bourriaud, 1998 quoted in Bishop, 2006, p 165]

But this is not without participation:

> "...an art form where the substrate is formed by inter-subjectivity, and which takes being-together as the central theme, the 'counter' between holder and picture, and the collective elaboration of meaning" [Bishop, 2006, p 167]

While the 1998 essay is quite recent in the context of this artform's history, Gustaf Almenberg goes one step further to consider "The Age of Participation", escalated by the internet in the 2000s [Almenberg, 2010, p 2], to consider how participatory art may grow and develop in this new technologically enabled scale.

3.2.2: Participation

In the novel "Notes on Participatory Art: Toward a Manifesto Differentiating it from Open Work, Interactive Art, and Relational Art" Gustaf Almenberg presents the history of participatory artwork along with parallel histories and obliquities from conceptual art and other related fields.

Of particular relevance and use for this thesis, Gustaf lays out 10 points of the manifesto:

1. Participatory Art is about exploring, within an aesthetic context, the many emotional facets of the creative moment as such and of one's own creativity, as opposed to solely contemplating the results of other people's creative moments and creativity
2. Participatory Art has its exploratory focus, not on line, color, or form per se, not on light or texture, not on narrative or representation, not on still life of the human figure or on a landscape, but on the very basis for all art: on creativity as such and on the feeling-thinking nature of creativity, as well as on what happens in the creative moment itself.
3. Participatory Art consists, technically speaking, of a number of parts/elements (whatever physical or electronic) which the "spectator" can (re)arrange, (re)assemble, or (re)combine into whatever whole that the spectator finds interesting -- aesthetically or otherwise
4. Participatory Art builds on the relations between its parts/elements/forms and on the changing of those relations on a trial and error basis to suit one's own satisfaction as a "spectator". Participatory Art also builds on the relationship between those parts and the choices made by an individual, thus involving intuition and individual responsibility
5. Participatory Art changes the role of the spectator dramatically -- from a relatively passive role of contemplating in front of an object and/or "completing the work of art in one's mind", to a distinctly more active role; not only in the sense that the spectator gets physically involved, but in particular also in that the spectator has to make strategic choices for the full work of art to come into being
6. Participatory Art retains the artist's task as intended "to make forms and colors living and capable of arousing emotions" (Mondrian 1937) while at the same time intending to give space to the creativity, physical involvement, and experience of the "spectator"

7. Participatory Art enables the expression by the activated “spectator” of a combination of intellectual expression (making choices), aesthetic desires, and a sensuous experience (in touching, picking up, and moving the various parts of the Participatory Art work).
8. Participatory Art probably most often manifests itself as some kind of abstract sculpture, but that is not the only option.
9. Participatory Art contains a unique measure of quality. A successful work of Participatory Art is judged also by the extent to which it makes space for the active choices made by the “spectator” for the work of art coming into being.
10. Participatory Art is founded on a vision of human creativity being intrinsically of great interest and so in itself worth pursuing for people in order to gain a deeper understanding of what it means to be a human being

[Almenberg, 2010, p 9-11] *note some shortening was done for clarity

Throughout the work, Almenberg proposes many characteristics of participation and how it departs from other modes of interacting with the artwork. The most notable characteristics that Almenberg analyzes that apply to this thesis are serendipity, choice, creation, expression, and the idea that participatory art is an art of many components.

Adding an opportunity for chance and serendipity is a key feature of participation that separates it from a fully curated interactive experience or performance. Duchamp first incorporated this as an early predecessor to participatory art through his paintings and readymades in the early 20th century [Almenberg, 2010, p 37], relinquishing the visual control of his end product.

Choice is another clear differentiation between participation and interaction. Within interaction, the choices are often fully set, but within participation, there is often more freedom, even if within boundaries.

The element of creation -- that the participant is actively participating and in so actively *creating* the artwork -- is perhaps the cornerstone of participatory artwork. And within this creation, there needs to be an element of freedom of expression and the expression of the participant. Almenberg argues throughout his analysis that the artist needs to enable this expression to be aesthetically pleasing.

It could be argued that perhaps the challenge of participatory artwork is to cohesively integrate these different expressions without losing their novelty. This of course relates to the last distinction aforementioned, that participatory art is one of many components.

From these considerations, a clear scope of the artist’s duties emerges. The role of the artist is to present a provocation and design a space and method for participants to respond to this provocation. And to do so in a manner that assembles the participants’ components into an artwork while preserving their essences.

In particular, this experience that the artist must build should be “neither so difficult to use that the other person shies away, nor so easy to use that there is no risk of “failure” and hence no

challenge.” [Almenberg, 2010, p 22]. This point is particularly pertinent to *A Counting,* in which the participant is expected to and feels obligated to count to 100 accurately in the language that they are serving as a representative of.

Participation enables an experience, and in some cases, a more nuanced, layered, and profound interaction with the artwork. This can in turn both investigate and emphasize the concepts that the artwork is developed and seek to develop. This is perhaps why participatory and conceptual artworks have many overlapping histories. This is where the role of the artist becomes important again, but also the intention. However, it still stands to question whether this intention will always affect the participatory experience.

3.2.3: Participation roles and results in *A Counting*

Understanding the definition and boundaries of participation, it is worthwhile to consider the roles of participants in *A Counting: Sign Language.* Understanding the definition and boundaries of participation, it is worthwhile to consider the roles of participants in *A Counting: Sign Language,* in which participants both record and transcribe counts in sign languages.

This extensive role space creates a unique genesis of a diverse community that is actively building an evolving artwork in a distributed and allochronic manner. Such a community can grow more rapidly thanks to the digital and asynchronous nature of the work -- anyone can participate at any time from anywhere and can pass links to their friends. Since participants receive unique links and can log into the website again, they can always return to transcribe new media in their languages, adding further continuity to this populace and work as it continues to grow.

This unique continuity, diversity, and scaling of the participation in this work are enabled by modern technology. As Almenberg points out in the preceding analysis, this comes hand in hand with the “age of participation”, a social understanding that participation is happening at a new scale by technology, with the internet as a “multiplier for creativity” [Almenberg, 2010, p 2].

When participating in *A Counting,* there is a cognitive load placed upon the participants -- counting or reviewing numbers 1 to 100. Many participatory artworks require a cognitive load, but this artwork also requires a setting where one can do this since the work is not in a dedicated space.

For the past year and a half due to COVID-19, the inequities of different personal spaces have been exposed, from online learning to difficulties in working and receiving social and medical care at home [Spievack, 2020]. Participating in this video portrait in particular is extremely indicative of this period of video calling and video communication, but also may add to this tension of individuals not having access to a dedicated place to do a task like this.

We use backgrounds in the portrait to expose this aspect of the diversity of the participants -- not just the linguistic stylings of signed languages, the hands themselves, but also where they sign. Different home settings, work settings, and more.

Figure: From the user test detailed later in this document, showing a participant counting in a mask in a public library. This type of background and time capsule of different computing conditions can only be done in an asynchronous artwork such as this one.

The participation of this artwork is not just the participant and their language, but also their environment. It is asking for a moment of their lives and a testament in the visual representation of their life settings in contrast with the language in which they live that life.

This is because, unlike a variety of participation options, part of the conceptual grounding of this project is that language does not occur in a vacuum. Language is an integral part of one's involvement in society and existence in their daily life, and as such it happens at the locus of multiple oppressions and distractions of the real world.

Legibility is the goal of transcribing these diverse samples in *A Counting*. But in asking for legibility, also leads to exclusion, curation, and potential bias in this aspect of participation within the project, as exemplified above. The task asks the participants to pass subjective judgment based on the legibility of the number of media presented to them. There are layers of potential bias introduced -- someone may decide that they do not wish a certain number to represent their language and this decision may not be based on the legibility of the number itself, but rather on their preferences around accent and aesthetic.

And while this participation structure opens potential for bias and exclusion, it relinquishes itself to the language community and all nuances (both equitable and inequitable) within it.

3.2.4: Capital and attribution

These actions of recording and transcribing can be quite engaging and require cognitive load but also a sense of cognitive capital at the population level that the entire *A Counting* series relies on. By its nature, the portrait needs a participant base that can transcribe or record in many languages. Furthermore, a population that will accurately do this task online.

Many artists engage with ideas of crowdsourcing in their works. In particular, from a critical stance that crowdsourcing can be used to exploit workers with digital tasks and little to no wages, such as Mechanical Turk [Semuels, 2018]. This critique of labor, ownership, and credit is not new, especially in the realm of participatory art, and is important to consider in the evaluation and definition of participation in this thesis.

One such example is Agnieszka Kurant, who considered an extreme of this type of crowdsourced labor in her “Assembly Line” work, which fused selfies of Mechanical Turk workers into a sculpture [Kurant]. Meanwhile, artist Aaron Koblin uses Mechanical Turk to collect sketches in his “Sheep Market”, of which he states:

> The inspiration for The Sheep Market project stems from the urge to cast a light on the human role of creativity expressed by workers in the system, while explicitly calling attention to the massive and insignificant role each plays as part of a whole. [Koblin].

A Counting: Sign Language, which uses crowdsourced methods, rejects the methods of Mechanical Turk and similar platforms. It does so by having its own technology to enable the artwork, crediting its participants, and not seeking to abstract contributions. Additionally, this participation is not just trading information or completing a task -- participants are actively creating the artwork together.

80 languages transcribed across 211 participants

Albanian 1
Antonio Giulietti
Amharic 2
bede
Ewunet T.
Arabic 5
Dima Alsakka
Holden M
Irene Siegel
Trevor Shealy
Arabic, Egyptian 2
Mariam Elnozahy
Arabic, Lebanese 1
Amaal Doukmak
Armenian 1
Bahasa Indonesia,
Indonesian 1
Gabriela O
Bengali 1
Shuprima Guha
Bosnian 1
Moira Spahić
Bulgarian 1
Marina Grozdanova
Cantonese 3
Charles Eppley
S R
Tania
English, African-American 3
Charles Eppley
Cynthia McCall-Torres
J Allen
English, American 7
alexis vondran
Aubrey (Simon) Simonson
Ayana Evans
Charles Eppley
Jeremy Karstens
Linda Krause
English, Australian 1
Charles Eppley
English, British 1
Shuprima Guha
English, United States 1
Farsi 3
Alejandro FrancoI
Delara Mohtasham
Trevor Shealy
French 7
Amoge Ezike
Indonesian 1
Gabriela O
Isoko 1
Shaj
Italian 3
Cynthia McCall-Torres
Rosie
Wilbur Vittorio Webb
Japanese 5
Justin
Miyo
Nich Kurtz
Uche
Korean 3
Amoge Ezike
Tiffany Lambert
Laotian 1
Latvian 1
Krista Roze
Lenape 1
Jim Rementer
Malay 1
Wei Han Lim
Maltese 1
Mandarin, Chinese 6
Ethan Nevidomsky
Evgenii Liakhov
Irina Vasilyeva
Mykhaylo Glukhov
Serbian 1
Shanghainese 1
Frank Yu
Slovak, Slovakia 1
Adelka Polak
Somali 22
Ally Hong
Andrew Clarke
Ann
Betsy U.
brenda
Charles Eppley
Cynthia McCall-Torres
Edwin Huitzil
Erika
Jess
Juliana Castro V.
Katherine Serrano
Maria
marilyn
Niama Safia SandyI
CMJ
Crystal Theaker
Devin Nau
Eden
Evan Fladager
Furqan Jawed
Jacqueline Gonzalez
Jennifer Bochenek
Jeremy
Jess
Justin Blinder
Katherine Serrano
Kerry MacNeil
Kofi Essel
KRB
Lorna Fossand
Matthew
Michelle Vu
Morgan O'Brien
Nina
Rachel L
Ri
S R
Sabrina
Sam Irwin

Figure: Credits for participants that recorded the content for the portrait and for those who transcribed languages on the website.

Along with this attribution and agency, *A Counting* seeks to build a distributed, linguistically diverse community that is contributing to a linguistic portrait together -- not a labor force.

3.2.5: Strengths, critiques, and challenges of participatory art

Fundamentally, one of the largest weaknesses and reasons for participatory art not taking off faster [Almenberg] is the art market issue presented in 3.1. Within participatory art, there is the additional issue of how one values an incomplete work without knowing exactly what the participation will shape the finished product into. Furthermore, there is a lack of a signatory artist to build a brand around if an artwork was not just made by that artist [Almenberg]. Sol Le Witt, despite his famed drawing instructions, had a variety of successes in solo signatory works that helped build momentum for his more conceptual and distributed artworks [Almenberg].

This is why the university backing and setting of *A Counting* is additionally important and why the editions of *A Counting* are done with presenting partners. This avenue of presenting partners for artworks makes this type of art more sustainable. *A Counting*'s web structure also

cuts down on the cost and resources of having participation locked into a gallery setting -- allowing for people to participate from anywhere at any time, adding another dimensionality towards the participatory process of the work.

Participatory artwork opens up opportunities for completely new creations and decisions of the artwork that the artist may not be able to envision for the work. However, some of these contributions may not fit in the artwork itself either due to aesthetic differences or conflict with other participants. There is thus a concept of an “aesthetic safety net” that Almenberg introduces throughout his notes.

Within the *A Counting* framework, this exists in that the work is built with software that can be modified at the artistic discretion of the laboratory but also through the filtering of transcription -- adding another layer of participation but also legibility checking for any numbers. All videos are also cropped and centered, with ones that have identifying characteristics or file corruptions being filtered out. This is a literal safety net poised to preserve consistency across the portrait but also to protect the contract of participation, as participants are consenting to the usage of their hand(s), not their faces or other identifying information other than what they give us such as their name for attribution.

But this chance and serendipity, the ability to expand the space of artwork away from the artist, is part of what empowers participatory artworks and evokes new possibilities and joys. From seeing different backgrounds to new languages to beautiful manicures -- there is much serendipity in the participation of this project. Almenberg claims that participatory art and its centering on the creative process through participants contributing to art can “express the great variety of emotions that enter into the creative movement” [Almenberg, 2010, p 168].

Claire Bishop interrogates the relationship between participatory art, history, and social relationships. Bishop implores artists and spectators to consider the relationship of labor and equality in this type of art: “...quality and equality, singular and collective authorship, and the ongoing struggle to find artistic equivalents for political positions…” [Bishop, 2012, p 3]. With respect not just to the artwork but to the exhibition of the work itself, Bishop has another layer of criticism:

> ...there is usually the objection that artists who end up exhibiting their work in galleries and museums compromise their projects’ social and political aspirations; the purer position is not to engage in the commercial field at all, even if this means losing audiences…it [gallery thought] also reinforces the hierarchies of elite culture. Even if art engages with ‘real people’, this art is ultimately produced for, and consumed by, a middle-class gallery audience and wealthy collectors. [Bishop, 2012, p 37]

This criticism is extraordinarily important for this project in considering the consequences and connotations of labor and politics in these works, especially at the velocity of digital participation. However, many participatory artworks, including this one, have a large public component to their practice, development, and exhibition. *A Counting* is exhibited in galleries, but also completely

online for anyone to view. And there is thus an act of resistance and protest in the ability of any participant to be represented in artwork in these spaces. It is a reclaiming of sorts, especially when one considers that the participants in this artwork are credited and minority languages are prioritized for display.

This is why the role of participation must be centered on creativity and adding to a piece. Good artistic decisions both bring in multiple veins of exploration, but also draw good boundaries. These boundaries, in the form of clearly defined and structured experience for participation and attribution when applicable, can set up the difference between bad and good participatory art. After all, a successful valuation metric for this type of art may be the experience of the participant -- not just their emotional response, since some artworks may need to elicit negative ones, but in how this participation enabled them the agency and creativity to shape the artwork.

This brings about issues regarding rigor in engagement and social disciplines regarding participatory art. Social and political issues served as the catalyst and framing for much participatory art, introducing a “politics of spectatorship” as Bishop claims throughout her works and criticisms of participatory artworks. There is a need for participatory artists to engage with methodologies from the social sciences that initiate and influence the participation and inception of their works, and in this engagement, Bishop claims that:

> Participatory art demands that we find new ways of analyzing art that are no longer linked solely to visuality, even though form remains a crucial vessel for communicating meaning. [Bishop, 2012, p 7]

Theorists like Almenberg find this an exciting and worthwhile endeavor, suggesting that this approach may need to be more nonlinear, but that this may better reflect the human spirit, condition, and creativity:

> Art, however, fundamentally deals with human nature and the relationship between human nature and the world it finds itself in...it is logical to reconnect with predecessors but at a different level, with a different perspective than they could possibly have taken. - [Almenberg, 2010, p 176]

Participatory art has not only the power and opportunity of an artist(s) to explore realities and concepts and histories, but also to expose and empower human creativity and response in the framing of their choosing. This exploration is powerful and can push the feedback system of culture and art to new heights.

3.3 Video as a medium

> ***"As with every technical medium, technological development marks the peculiarities in the way video aesthetics develop." [Spielmann, 2008, p 133]***

What does it mean to make participatory web artwork with video as a medium? What dynamics of power and representation do visual, particularly video media, entail? How does video as a medium further the conceptual grounding of this portrait of signed languages? These are just some questions in this arena of inquiry regarding a rich area of generative video artworks that *A Counting: Sign Language*, by definition, engages with.

It is important to consider this is primarily by definition -- *A Counting: Sign Language* draws more from the participation and conceptual groundings as an artwork -- it is not meant to be evaluated as video portraiture for the sake of video portraiture. *A Counting: Sign Language* recognizes video as a medium -- a tool towards conceptual artwork.

Many conceptual artists have engaged with video in the past, including artists such as Bruce Nauman, who utilized video to experiment with notions of body and time [Frohne]. Nauman had works such as "Lip Sync" in 1969 which featured a video of part of his face with an out of sync audio track of "lip sync" [Forhne]. In the 1990s artist Gary Hill experimented with reading and video in his "Remarks on Color", featuring a child reading from Wittgenstein's Remarks on Color, showing the juxtaposition of a child reading to very intricate philosophical wordings [Hill, 1994]. In these cases, video is used to showcase a juxtaposition or division from the audio of the video, while *A Counting: Sign Language* uses video and the editing and centering of hands in video as a unification between different signers and one count.

Along with the end video, the experience of recording holds weight in the concepts of accessibility and allowing signed languages to hold space in this portrait. It also affects the final work. Its methodology of recording video feeds into the video portrait -- creating a unique reflexive nature within the artwork. Yvonne Spielmann in *Video: The Reflexive Medium* argues that video is not merely an intermediate stage between analog and digital, or even a technology, but a medium in its own right that both responds to and feeds cultural aesthetics [Spielmann, 2008, p 130]. This evolution from tool to medium brings about an important discussion of this type of work, as it is enabled by technological and video tooling.

Recordings and artworks like this often happen in live spaces with filtering and abstraction, such as the gallery scale video and code artworks of artists like Zach Lieberman, who often has interactive components where users see themselves in the artwork, but in an abstracted way and in a live feed that does not stay after they are gone.

This is part of the nomenclature of "video portraits" for this work. A portrait is often thought of as an artist making a likeness of a subject in a direct way through observation then creation through a medium. *A Counting: Sign Language* both removes the artist from the loop and also reconstructs the subject from the individual signers into a new one -- different hands in a count of 1-100, or more abstractly of the United States and its linguistic diversity. This construction is

done on a technical level by digital video technology but also visually by reframing the original footage of individuals into close ups of the hands in sequence -- similar to purposefully framing any portrait.

This is all possible from the participation of users who are willing to record themselves for the artwork. Using the technology in this thesis and through the experience of participating in the work, the participant becomes the videographer and actor in a growing cinematic universe of the portrait. The backgrounds, the signer -- these are all added into the same portrait.

Furthermore, within this project, there is also a purposeful divorcing from audio. Much video art involves sound. However, for this artwork, we do not even collect audio during the streaming. This is partially an engineering decision, as it makes the files smaller, but also an artistic one. This video portrait is about holding space for signed languages -- and while the backgrounds of the participants are partially visible, removing any possible sound reserves the space for the language and allows the viewer to experience the language as it is intended -- fully visually.

3.4: Public participatory artwork

> ***Public art is a reflection of how we see the world – the artist's response to our time and place combined with our own sense of who we are. - Association for Public Art [Association for Public Art]***

An important component of the conceptual framing and participatory component of *A Counting* is that it is a public artwork -- anyone can access all components of the artwork -- from the creation to the portraits themselves.

3.4.1: Public art

The common association for many regarding public artwork is large sculpture works in public places, which are a component of such artwork. However, public art is agnostic regarding medium and execution:

> What distinguishes public art is the unique association of how it is made, where it is, and what it means. Public art can express community values, enhance our environment, transform a landscape, heighten our awareness, or question our assumptions. Placed in public sites, this art is there for everyone, a form of collective community expression. [Association for Public Art]

In the case of *A Counting,* the public site of the artwork is the website itself. It is a publicly accessible statement and meditation on a united and contemporary portrait of language that is motivated and shaped by participation.

3.4.2: Public participatory art practice

> ***"All stories emerge in the midst of complex and uneven relationships of power, prompting certain questions about production: Who tells the stories and who***

> ***doesn't? To whom are they told and under what circumstances?" - Kay Schaffler and Sionie Smith, Human Rights and Narrated Lives, 2004***

A Counting is a public participatory web artwork. As a fully publicly accessible artwork on the web and in exhibition spaces, it continues a tradition of web art that blends the lines between gallery and web.

In particular, it pushes in the direction of artworks "going public", a concept analyzed at length in the same titular novel, *Going Public: The Art of Participatory Practice* by Elizabeth Miller, Edward Little, and Steven High. Works that "go public" are responding to a need for artwork to engage with the populus, answering the urgent question of what it means to actively create with and for the public. This urgency within the participation of works like *A Counting* is that they necessitate the space for participants to engage with and add to the conversation of pressing present issues.

Miller et al argue there is a complex system within the creation, reception, and critique of public and participatory art that can be evaluated through research and artistic practices to address and avoid unintentional harm such as appropriation and exploitation. This is examined with an understanding that participation and working with or for the public now involves technology at scale.

In the modern age of "data sharing" [Miller et al, 2017, p 110], it is important to create standards that do not invisibalize intellectual labor and that prevent exploitation of data generated by these types of artworks [Miller et al, 2017, p 110]. Consequently, it is important to bring value to more qualitative datasets and histories such as oral histories and how they are curated, stored, and engaged with [Miller et al, 2017, p 111].

> "Oral sources are not objective; they are highly subjective. That is their strength. Oral history tells us "not just what people did, but what they intended to do, what they believed they were doing, and what they now think they did." [Alessandro Portelli, folklorist, 2017, quoted in Miller, et al, p 111]

In these contexts, "data" is a large, organic definition -- where so many artifacts, such as oral history, public perception, and language itself, are seen and described as data. Driven by our cultural perception of the need to quantify and qualify the human experience and history, it is important to also hold these datasets to their human and artistic values, as precedents and cultures, not just quantifiable entries.

This infrastructure around the sense of data of all types begets a new question of curation and development of public artworks -- after all, if this type of data is subjective by the human experience and systems, so is both our interpretation and utilization of it in the artwork. There is thus a responsibility to engage the public with this data in a responsible way and extrapolate it and reinterpret it within the artwork to prevent unintended harm.

This is part of the reason that *A Counting* does not modify the video footage through interpretive methods such as special effects or filters. All participants are accredited by chosen names on the website and in the portraits. And why so much consideration was taken into the technical security and cropping out any identifiable video footage and metadata. *A Counting* goes public both in terms of participation but also in active co-creation, placing language identities and history as artifacts of public and social creation, rather than a removed census. This approach is a public and human-centered way to reimagine and enable these pieces of culture through art.

Within their theoretical framework around artworks "going public", Miller et al also propose the following principles regarding public artworks and public art practice:

1. Art, story-sharing, and public engagement are essential to strengthening democratic processes;
2. A shared commitment to collaboration and shared authorship strengthens the work we do;
3. Opening up otherwise closed research and artistic processes fundamentally changes the relationship to source or host community, thus mitigating some of the recurring political and ethical problems of voyeurism and appropriation
4. Collaboration and public engagement place ethics at the heart of our research and creative practice; and
5. Working across differences over time generates insight that can affect social and political change
   [Miller, et al, 2017, p 263]

These principles are particularly salient for *A Counting,* and as the development of this project and document suggests, the utmost goal of this work is to generate spaces and representations for marginalized languages through social and public practice methods in a participatory artwork.

3.5: Social practice

A Counting is motivated by the social justice issue of the underrepresentation and further marginalization of intersectionally oppressed individuals by vehicle of the Census. As such, the way the artwork is developed and grown is done in a socially conscious manner.

While much of this is out of the scope of this thesis, it is important to consider the precedents of social practice and socially engaged artworks and art collectives. There are several artists, historians, and critics in the realm of socially engaged art and social practice. Social practice, for this thesis, encompasses artwork and artists that utilize human discourse and interaction as a medium regarding societal causes through artistic expression, communication, and thinking.

One such model for the *A Counting* series and similar works is represented in the work of artist and activist Hank Willis Thomas. Through projects such as *For Freedoms, Cause Collective,* and *Question Bridge*, Thomas and his team have developed methodologies to tell multimedia stories while engaging with marginalized communities [Arsty.net].

> *In the wake of the 2018 midterm elections, For Freedoms: Where Do We Go From Here? explores the role of art and visual representation in American civic life through the work of the For Freedoms collective.* [icp.org]

The works above, as well as others in Thomas's portfolio, encompass many individuals lived experiences of different identities at the edge of current, urgent events. In particular how these experiences show the role that art can play in life.

Another artist whose work exists within the social practice is Tania Bruguera, who developed *Arte Útil,* which encompasses a methodology and portfolio of work and programming [Bruguera]. As part of the methodology, the goals of each Arte Útil project are outlined below:

1) Propose new uses for art within society
2) Use artistic thinking to challenge the field within which it operates
3) Respond to current urgencies
4) Operate on a 1:1 scale
5) Replace authors with initiators and spectators with users
6) Have practical, beneficial outcomes for its users
7) Pursue sustainability
8) Re-establish aesthetics as a system of transformation
[Bruguera, arte-util.org]

In particular, point 5 speaks to the purposeful reframing of spectators to users and participants, something very centralized in the *A Counting* series as the artwork becomes centered around the experience and interaction with the community it creates or seeks to interface with.

The above examples are just some of the many social practice-centered artworks that seek to leverage art as a nexus of conversation and civic engagement across a variety of individuals with current and urgent issues. *A Counting,* with its agency of representation of the commonality of language through a multimedia portrait, along with grounding within issues of what happens when marginalized groups aren't counted, stands within this realm.

It is important to note, that while *A Counting* does center around justice issues, it is within the MIT Media Lab, which in itself, at the highest level, functions as a corporation [MIT.edu]. MIT has historically released research that has been harmful towards marginalized communities, such as projects that perpetuated visual and urban bias through the use of problematic research methodologies like Mechanical Turk [Hidalgo, 2016].

Therefore, while this thesis and the work it exists within is not a technology project, but rather an artwork enabled by technology, it is important to consider efforts at MIT that seek the advancement of justice and representation as it intersects with media and technology. There are several methodologies of design, engineering, and planning that have emerged that actively engage with societal issues and communities within MIT. Networks and practices such as Design Justice [Costanza] and the Algorithmic Justice League [Buolamwini], seek to explore

how art and design can be used to combat and understand biases and social issues. *A Counting* and works like it seek to show that not only can socially conscious and practicing engineering and design occur at MIT but so can art.

3.6: Artwork within disability

> ***"...but rather what needs to change are the problematic, limiting social construction and perceptions of disability in culture. What must change for progress to happen, and for our visual history to be a part of such a movement, are often our acts of viewing and interpreting" [Millett-Gallant, 2012, p 6-7]***

In the United States, the federal government defines disability as a legal concept under the Americans with Disabilities Act (ADA), passed in 1990 [ADA]. It is important to remember that in the context of the ADA, "disability" is a legal term rather than a medical one. Therefore, it entitles those under its definition to legal protections against discrimination and gives them the right to accommodations [ADA].

However, despite this focal legislation, disability is a social model, as argued by disability studies and by activists. This is an important distinction because within a social model comes structured and systemic discrimination against those with disabilities. This manifests in all aspects of life as those with disabilities try to lead their lives in society.

While this project is not explicitly centered around disability, within the context of today's spoken language dominated world, signed languages are used by those who are disabled under this social model and it is impossible to structure an analysis around the discrimination and marginalization of signed languages without considering those of disability.

3.6.1: Disability studies and ableism

Disability is not based on one individual's circumstance, but rather how society is structured -- if all spaces accommodated all needs, no one would need disability accommodations. This is the cornerstone of a **social model of disability**, which includes a very distinct and important difference between disability and impairment:

> The social model of disability, by contrast, distinguishes between impairment and disability, identifying the latter as a disadvantage that stems from a lack of fit between a body and its social environment.
>
> Oliver notes that "impairment is, in fact, nothing less than a description of the physical body." ([8], p. 35) Disability, by contrast, is the "disadvantage or restriction of activity caused by a contemporary social organization which takes no or little account of people who have physical impairments and thus excludes them from participation in the mainstream of social activities" ([8], p. 22).
>
> [Goering, 2015]

With these definitions and models, it is possible to understand by societal mechanisms that a deaf or hard of hearing individual becomes disabled in a society dominated by spoken English, as they are not accommodated.

Additionally, while “hearing impaired” is a medical term and by this model, this could be the **impairment** that characterizes deafness and hard of hearing, it is important to note that this is not the language that should be used to refer to Deaf individuals. This medical language is exactly what this social model asks to be distanced from. Additionally, “hearing impaired”, while a medical condition, is not inclusive to all that experience deafness or are hard of hearing. [CAD]. This nuanced approach to language is incredibly important in understanding and providing suitable accommodations for everyone who is disabled in society.

Such oversights in accommodations and implicit and explicit social discrimination are examples of ableism:

> Ableism is discrimination and social prejudice against people with disabilities and/or people who are perceived to be disabled. Ableism characterizes people who are defined by their disabilities as inferior to the non-disabled [Linton, 1998]

Ableism, like all social models of discrimination, integrates its way into every aspect of daily life. Many disability groups have released several legal and social briefings regarding the past and present dangers of ableism. Presently, the Communication Service for the Deaf (CSD) in the United States explains that 70% of deaf people are unemployed or underemployed [Waterfield, Newsweek, 2019]. Historically, ableism has been used to take away the rights of disabled individuals through abusive institutions with no regulations, restricting their rights to marriage and financial freedom, and even mass murders of disabled individuals [Siebers]. This vicious history against disabled individuals has contributed to an incredibly ableist society at large that has continued to oppress disabled people.

Disability studies is the field that examines this myriad of social oppressive models and histories that disabled individuals and groups face:

> “Disability studies has emerged in the last quarter of the twentieth century as a discipline that unites political activism, sociology, critical cultural analyses, and creative production...disability studies strives to establish a social model for disabled people as an oppressed group and forges alliances with theories and positions of other socially marginalized groups”
>
> “Ableism draws parallels to sexism, racism, classism….to encompass multidimensional practices (networks of social policies, attitude, and daily social exchanges) that marginalize disabled people...a social model that counters that ideology, policy, and attitude needs to change to accommodate, service, and provide equality for disabled people.”

> [Millett-Gallant, 2012, p 6-7]

Audism is a particular type of ableist discrimination that affects those with hearing impairments, centered around the belief that audible and spoken language is superior to sign languages and that deaf or hard of hearing individuals should conform to the hearing world through methods such as spoken language and cochlear implants [Berke, 2020].

> Audism is a term used to describe a negative attitude toward deaf or hard of hearing people. It is typically thought of as a form of discrimination, prejudice, or a general lack of willingness to accommodate those who cannot hear. Those who hold these viewpoints are called audists and the oppressive attitudes can take on a variety of forms. [Berke, 2020]

Along with deafness and hard of hearing, other individuals use signed languages in the United States. In particular, many individuals with autism use sign language as a method of communication [Bonvillian, et al, 1981]. Other nonverbal individuals of many backgrounds and circumstances utilize both formal and informal sign languages as a way to communicate [NAD]. Furthermore, many deaf individuals are “Deaf plus”, a term for deaf individuals who have other disabilities along with deafness [NAD].

Overall, it is impossible to have artwork around signed languages in the United States without engaging with the role of disability in history and art and vice versa.

3.6.2: Disability aesthetics

> ***Aesthetics is the human science most concerned with invitations to think and feel otherwise about our own influence, interests, and imagination.***
> ***[Siebers, 2010, p 25]***

Proposed in the titular novel, “Disability Aesthetics” by Tobin Siebers, this aesthetics field has to do with the perception and interaction of disability with historical and modern aesthetics. In particular, the role that disability and disabled artists play within modern art and its history and how the relational component of aesthetics itself has played a role in both the disqualification of disability in society.

Siebers defines **disability aesthetics** as “a critical concept that seeks to emphasize the presence of disability in the tradition of aesthetic representation.” [Siebers, 2010, p 2]. This is a unique framing, particularly concerning **representation**, which is a concept that this artwork embodies. **Representation** is a concept not only of abstraction -- using a symbol to represent something, but also representation in a contemporary aesthetics context of representation of identities in artwork and media.

This is acknowledged two-fold in Sieber’s analysis -- both of how disability is used in artwork to represent various concepts but also in how:

> ...under the pressure of the American culture wars, how the aesthetic representation of bodies -- individual and collective, organic and artificial--leads to the oppression of people with disabilities [Siebers, 2010, p 79]

This representation has real-world implications and consequences, especially considering in many ways it can “enfreak” individuals [Siebrs, 2010, p 21], distancing them from personhood and therefore seek to establish them as inferior. Siebers examines this in the context of **disqualification** “a symbolic process that removes individuals from the ranks of quality human beings, putting them at risk of unequal treatment, bodily harm, and death” [Siebers, 2010, p 23]. Through the analysis of the aesthetics role at mass scale, Siebers lays out how this disqualification has gone hand in hand with visual representation of disabled people and how it has led to harmful and, in some historical cases, deadly discrimination.

This particular **disqualification** speaks yet again to a core question of this project: Who counts? Which is easily interpreted to: Who matters? And in Sieber’s case: Who is a quality human being worthy of fair treatment?

This framework proposed by Siebers is invaluable in understanding the very real role that aesthetics and works of art and their treatment of subjects and concepts that fall into the social model of disability plays and how this power can create harm. By this logic, though, there is the hope that aesthetics and art can have a role in pushing for a more just society.

3.6.3: Relevant precedents

> ***“As described by art theorist Henry Sayre (1989), contemporary art, particularly performance and photography, refuse to be contained to any single context, viewer perception, or conventional attribution of meaning. Sayre uses the phrase “exceeds the frame” to describe such works’ projection of meanings that reach beyond the image itself -- beyond language, facts, and narrative -- and enter into viewer’s subjective, interpretive space…..between the symbolism, corporeal materiality, mediation, and lived experiences of the body on display.” [Millett-Gallant, 2012, p 20]***

For the context of this thesis, not all artwork that engages with disability is relevant. Especially as much artwork discussed within disability is often visual portraiture, photography, or sculpture [Millett-Gallant, 20]. But it is important to understand the categories many precedents fall into as part of this scope, instead of an in-depth history of all these precedents. In general, two themes emerge and are popularized in present-day art and art salient to understand the importance of nomenclature and framing of these types of works.

Disability art is a type of artwork about disability experiences and culture:

> The development of disability arts began in the 1970s / 80s as a result of the new political activism of the disabled peoples' movement. Previously, most art about disability was made by non-disabled people. When disabled people began using the arts to reflect their own experiences, this provided a very different perspective.
>
> [Sutherland, 2020]

Some definitions argue that artworks like “Alison Lapper Pregnant”, by sculptor Marc Quinn of fellow artist Alison Lapper, are disability art [Bragg, 2007]. However, for the purposes of this document the definition of **disability art** will be limited to artwork made by disabled people about the experience of disability. As such, artwork by Alison Lapper about her disability would be **disability art,** but not this particular sculpture, as Quinn is not a person with a disability.

Many artworks by non-disabled artists have a context around disability and accessibility for disabled spectators, but these are not disability art.

> The term 'arts and disability is used to describe arts projects specially set up for disabled people, but usually led by non-disabled people. [Sutherland, 2020]

This above is the framing of a project like *A Counting: Sign Language* -- it is not disability art, but is a participatory art project that is accessible and centered around signed languages, which are primarily used by individuals in various communities, most specifically and most prominently within disability settings and communities [NAD]. It is impossible, as argued in this thesis, to have a work about sign language without acknowledging ableism and disability art. Christine Sun Kim, whose work is talked about at great length in this thesis, is an example of an artist who engages with the social role of disability and inaccessibility for Deaf individuals in an audio-centered world.

James Castle, born deaf in 1899 in a rural mountain community in Idaho, was a self-taught artist whose work in mixed media drawings focused on nuanced and detailed views of life in Idaho [Castle]. Castle never learned formal sign language (it is said the family had a set of signs they used with one another) or read written English, but some of his artwork depicted language [Castle]:

> “Castle was also clearly passionate about language and communication, and created a large number of works investigating letters, signs, and symbols assembled into dizzying and ever-changing combinations. These sometimes filled books Castle bound from the same paper and packaging materials scavenged from around the property.” [Castle]

Castle is someone whose artwork stands as a testament to language deprivation -- his pictures are not about his life as a Deaf person nor any relationship with sign. But rather daily life in his home and the things that permeated it. Much of this style of work speaks to popular print media of his time and his interpretation of it, particularly ephemera [Castle]. In some manner, Castle’s work with language symbolics are linguistic portraits in themselves -- a meditation of written

English from someone who never learned it or a language of that same system but was constantly surrounded by it and affected by it.

There are several other examples of **disability art** that scope out different experiences, reflections, and provocations of living with different disabilities in society. Likewise, there are several artworks made by non-disabled artists that engage with disability, and these necessitate additional critical frameworks.

3.6.4: Critiques and challenge of artwork around disability

“Nothing for us without us”, a phrase made popular by the 1990s disabilities activist movement [WHO]. This shows volumes about how marginalized groups face work thrust upon them with wrong assumptions of their needs and priorities.

It is important to understand that this project is not meant to offer any type of solution regarding the use of sign language in any individual or systemic way. Rather, it is meant to provide an option for users of signed languages to partake in an ongoing artwork about language diversity in the United States. However, this project does engage with a marginalized language system and with issues of disability, while being situated at an elite institute (MIT) as a participatory artwork (and thus requiring the participation of these individuals to exist).

Towards the first point, MIT has produced work as an institute around disability without centering relevant communities, in particular concerning technology-focused projects. Projects out of MIT such as AlterEgo [Media Lab], a VR simulation of disabilities [Couch, 2016], and signing gloves [Reddit] have all faced criticism in various disability (in this case particularly Deaf) communities for their lack of involvement of their community and for creating solutions that were not requested. These projects are ableist in that they make assumptions regarding the needs and desires of individuals instead of engaging with them. These researchers also received extensive press and credit for these projects, often in a beneficial framing for them and their research.

Logically, this type of work often comes off as disingenuous or insensitive at best and exploitative at worst to members of these communities. It makes sense then why some individuals would be confused and trepidation regarding a project like this being within the MIT Media Lab.

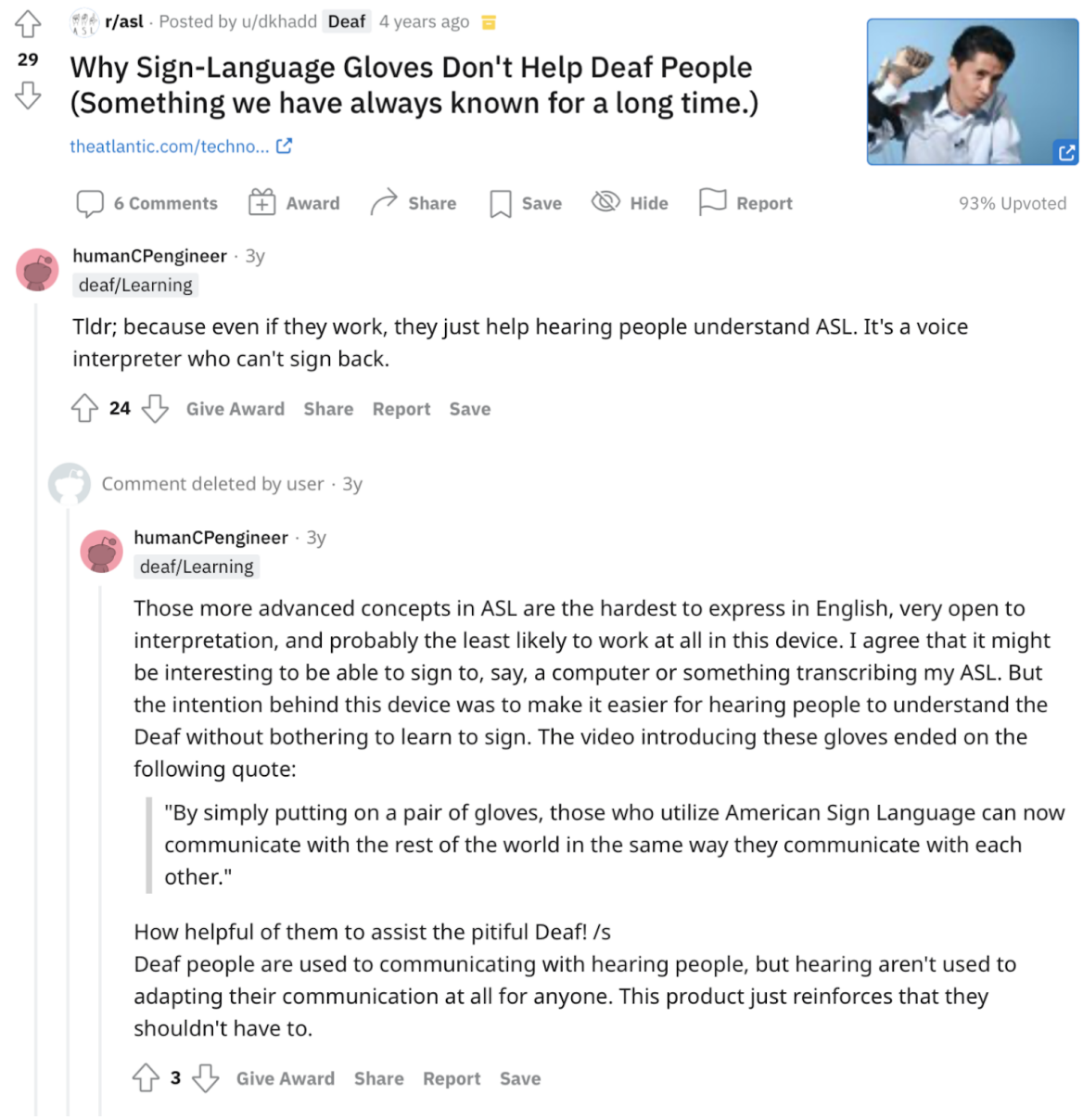


Figure: Examples of coverage and Reddit responses to a signing glove

Indeed, throughout my time doing this project, due to my background and situation at MIT many assumed it was a project regarding recognition or “solutions” regarding sign language, instead of an artwork to hold space for these languages. Signed languages are complex and beautiful languages that do not necessitate “solutions” and interventions but rather the respect and space that audio languages receive in our society.

For example, to be a hearing computer scientist to be rattling off about hand recognition for automated interpretation, considering that sign languages rely on the face and body, is not only ignorant but offensive to how it simplifies the complexity of signed languages. While there is a time and space for improving technologies and methods and resources for things like

interpretations so that more users of signed languages can receive the services and accessibility they deserve, such efforts necessitate involvement of those individuals throughout the process and should be open to criticism regarding intention and possible misuse.

Elizabeth Sweeney, art critic and professor, proposes an outline for such critique, in a manner that feels appropriate to consider *A Counting: Sign Language.* Sweeney asks the following questions of artworks that engage with disability in any way:

1. Who benefits? Career, media attention, ego, reputation, financially et cetera.
2. Who is in control of the message and the artform?
3. Will disabled people benefit directly or indirectly from being involved in this work and by viewing it or is this about enlightening nondisabled people?
4. What is the power relationship between the subject, the artist, and the audience?
5. What is the contribution of the disabled artist in the work? Passive or active? Do they have control of the project and how are they being presented?
6. Does this work challenge current assumptions and stereotypes about disability or does it reinforce them? If it is the latter, reconsider the question and benefits.
7. Have assumptions about disability and/or disabled people been made from a nondisabled person about the direction of this piece?
8. Are disabled people treated as artists or subjects, are they named and recognized as individual contributors? If they are not, reconsider questions 1, 3, and 4.

[Panel, 2020, Creative Capital, " Access in Content and Form", Kenny Fries and Alison O'Daniel]

3.7: Language, numbers, and art

> ***"What are the theories of change we embody in our own language? Language is the interface through which we engage complex systems." - [Nora N Khan, "Second Readings", 2020]***

A Counting, at its core, is about language. Throughout this thesis, arguments have been made regarding language as identity, language as justice, and how the concepts of language and counting are related and conveyed in *A Counting.* This section will examine existing work at the junction of language and art as well as art criticism and theory. In particular considering language within complex systems, language as cultural data with respect to the digital humanities, language in conceptual art, and lastly the treatment and involvement of alphabets and numerical systems as components of language in the artwork.

3.7.1: Language in systems

Infrastructure and social systems are fundamental in the pursuit of any justice and is something that is particularly a pain point in the context of accessibility as much infrastructure is built with ableist designs -- including language and metaphors within modern language [Khan, 2020]. In her essay "Second Readings" critic Nora N Khan reflects on language and systems --

particularly the limits of language. Language as a system both connects and alienates, both empowers and is weaponized. It is a system within a system to describe an infrastructure that logically we are all in but experience very differently. Language can limit us in expression and imagination of the possibilities -- perhaps further keeping us in the system limits that language itself was designed in.

One matter in which this occurs is at the site of labor of language acquisition and preservation -- there is a system to learn and preserve languages as cultures and peoples. Many signers have to seek out signed language services and education -- this is labor that goes unseen and unrecognized. In some ways, the preservation and use of signed languages is an act of resistance and demanding space. Signers also need physical space, not just digital and representational ones presented in this project. Signed languages are extraordinarily three-dimensional and spatial. In “What Can a Body Do?” Sara Hendren interviews the architects of Gallaudet University:

> Gallaudet’s architecture emphasizes the particular capacities and assets of deafness...Deaf Space isn't a plea to “make room” for deafness. It’s an unapologetic and joyous expression of the integrity and beauty of deaf experience, codified in a series of strong principles that inform the way the rooms here look and operate and feel. [Hendron, 2020, p 99]

Instead of this physical infrastructure, *A Counting* seeks to grant digital and representational space to marginalized languages as a statement against unjust infrastructures that promote language injustices. This is just one manifestation of sign language but adds a layer of permanence in making digital archives of different signers in one count across the nation.

3.7.2: Language as cultural data and identity

Such digital natures of *A Counting* begets a relationship to digital humanities, as this artwork is enabled by the digital. Digital humanities seek to enable and discover new frontiers of humanities studies with technology, intersecting with language at a level and framing that is particularly interesting for *A Counting: Sign Language.* In particular, with its systematic approach to understanding, cataloging, and analyzing knowledge from the social sciences.

In the novel *Cultural Analytics*, Lev Manovich seeks out better representations for cultural data -- including language. Manovich discusses at length the relationship between language, art, and representation of information and culture:

> Natural language was the only mechanism in humanities for describing all aspects of culture until the recent emergence of digital humanities. [Manovich, 2020, p 154]

Manovich, whose focus is on new media theory and visual cultures, also states the following as limits to what language can convey:

> Natural languages cannot represent small differences in analog dimensions that define aesthetic artifacts and experiences such as color, texture, transparency, types of

> surfaces and finishes, visual and temporal rhythms...In contrast, our senses capture such differences quite well. [Manovich, 2020, p 154]

This is a framing that is incorrect regarding signed languages in particular, as the artworks of Christine Sun Kim often show an understanding of sonic properties and relationship with sound. Signed languages are often described as 3D languages [Hendren, 101] and portions of these languages involve the physical properties of space and sonics (e.g. banging on a table to get someone's attention) [Hendren, 2020, p 100].

Later in the same chapter, Manovich goes on to explain that when considered in the context of how human perception makes art enjoyable and meaningful that "human senses and natural languages are complementary systems" [Manovich, 2020, p 157].

Therefore, it stands to reason that language and its histories can hold much more data than digital humanists may give credence to. This speaks vastly to the biases within how language is represented and considered in art and art histories. This is particularly true for signed languages, which have been left out of much of discourse.

Additionally, this is why the departure from language as data is important in this and other precedents, including the ones mentioned above. Language grounded in representation and identity is a more productive conversation for the spaces that many of these works aim to create, especially *A Counting* which is not positioned as a data artwork.

3.7.3: Alphabets and number systems in art

Many artists consider components of language instead of whole words or sentences to play to their concepts. This is what *A Counting* does with the idea of counting 1-100.

While most language systems have numbering systems, it is important to note that some languages do not have formal numbering systems [Everett, 1988]. One such is the Pirahã language, which only has words for small or large quantities, but not formal numeric digits [Everett, 1988].

Like an alphabet, numbers can never hope to encompass a language and can be misleading. Counting is a building block that everyone learns early in language acquisition. This association brings about a universal experience within a diverse representation of different languages and provides for a common ground for combination. Within the count, one can also consider the order critically and add randomness such that languages are not clustered in a notion of hierarchy and similarity, but serendipitously allowing each iteration of the count to present a unique permutation of voices or hands.
Numbers and alphabetical systems make up these building blocks of languages and can therefore construct powerful concepts around language and language identities.

Artists such as Xu Bing, who is famous for several mixed-media installations regarding the written Chinese language [Bing], play with the idea of written alphabets. Xu Bing's artwork "asks

viewers to consider how our cultural backgrounds, especially those shaped by language, fundamentally color our worldviews" [Bing]. His *Square Word Calligraphy,* a system for writing English with calligraphic brushstrokes nesting alphabetic characters into square (e.g. "Chinese") ideogrammatic forms, hones in on the viewer's expectations of language in its representation and our assumptions about it.

A variety of alphabet artworks have been designed by other artists. Derek Beaulieu's 26 alphabets for Sol Le Witt solicits 26 different artists to design alphabet based works in homage to Le Witt's process of having other artists interpret and create from his concepts [Beaulieu, 2016] and in Beaulieu's case, showing different interpretations of the written English alphabet and concepts underlying it.

On a more anthropological angle, Frédéric Bruly-Bouabré's Une Méthodologie de la nouvelle écriture africaine, "Bété;" L'Alphabet de l'ouest africain (A Methodology of the New African Writing, " Bété," Alphabet of West Africa) featured a linguistically motivated ethnography that resulted in Bruly-Bouabré formulating and documenting a writing system for the Bété people. [Tobias, 2012].

While there were more precedents of conceptual works around alphabet systems, there are a variety of artists that engage with counting and numbers, particularly in the infinite nature of numbers and even counting itself. One particular example is Polish painter Roman Opałka, who explored the idea of progression and infinity within his numerical and color paintings around counting:

> It was then that the principle of harmony and permanent systematization – that he had begun in 1965 – started to govern Opałka's art, taking the form of the "concept of progressive counting". [culture.pl]

Opałka even recorded audio of himself counting with photos of his face registering these counts, creating art through the act of counting both auditorily, internally, and through painting [culture.pl]. This auditory nature of counting is not just seen in Opalka's art, but in other conceptual artworks. One such is On Kawara's *One Million Years*, a conceptual piece in which alternating performers read past and future years -- a performance of a count of time with the human voice [Steinhauer, 2015, 2014].

This is perhaps the closest language artwork towards *A Counting* -- a meditation on the act of counting and what concepts can be derived from it, in multiple media and through an experience in itself.

3.8: Signed language histories, media, and cultures

> ***'...Were anyone to use a language addressed not to the ear, but to the eye..the present inferiority of the deaf would entirely vanish' -[John Burnett, 1835, Tales of the Deaf and Dumb quoted by Hendron, 2020, p 106]***

The history of signed languages is a unique and intersectional one. With the focus on the United States, as this thesis seeks to build a “gestural portrait of the United States” this section seeks to add a brief historical context to signed languages in the United States as well as social and artistic framings and precedents for the treatment of signed languages.

It is important to note that while this section focuses on the United States, and indigenous signed languages and American Sign Language (ASL), that any signed language is welcome in *A Counting* as there are a variety of immigrants and cross-cultural experiences in the United States.

3.8.1: History of signed languages in the United States

The history of sign language in the United States is a two-fold one: firstly, of the contact and eradication of indigenous peoples, and with them endangering or eliminating their signed languages; secondly of the development of American Sign Language (ASL) from LSF (French Sign) and community English sign languages of colonists in the 17th and 18th centuries.

Before colonization by Europe, indigenous communities had several distinct sign language families with a variety of tribal and regional dialects that have been documented through linguistic and oral history, as well as anthropological study in the United States [Davis, 2017]. There are two broad categories and use cases of indigenous sign languages: languages that emerged from one or more families for deaf individuals to communicate in their tribe, and languages that emerged for communicating across tribes and with traders [Kelley, 2003].

Plains Indian Sign Language (PISL), is perhaps the largest and most documented sign language family in the indigenous peoples of North America [Davis, 2017], having replaced Plateau Sign Language when the Crow Nation brought it north before colonization by Europe [Davis, 2017]. It was primarily used to communicate across language boundaries in tribes, and secondarily for those who were deaf or hard of hearing, who often had their own local signs for more expanded grammar [Davis, 2017]. This shows an entirely different language paradigm than much of European history -- utilizing a common, gestural language instead of one of the groups’ languages dominating.

In the Southwest, the Navajo people followed the family sign languages model, where sign languages were localized to families and tribes to communicate with deaf and hard of hearing members, not motivated by communication across tribes [Davis, 2017]. Keresan Pueblo Indian Sign Language (KPISL), also developed in tribes from one family in what is now New Mexico, went on to become the first language for all deaf members of the tribe of its origin and many surrounding tribes [Kelley, 2003], showing a formalization of extensive signed languages and their spread. KPISL, unfortunately, is extremely endangered in the present day [Kelley, 2003], due to the colonization of the tribes that used it.

Along with linguistic history and communication, many tribes utilized signed languages as methods for storytelling and rituals. There are countless other indigenous sign languages, many of which are exceedingly rare and endangered, and some of which influence modern American

Sign Language (ASL). However, many tribes are still practicing their signs today and fighting to preserve them and their traditions:

> Remarkably, members of these Indian nations and others are still using and maintaining the traditional Native American ways of signing for a broad range of discourse purposes, including traditional storytelling, rituals, legends, prayers, jokes, games, conversations, and personal narratives [Davis, 2015]

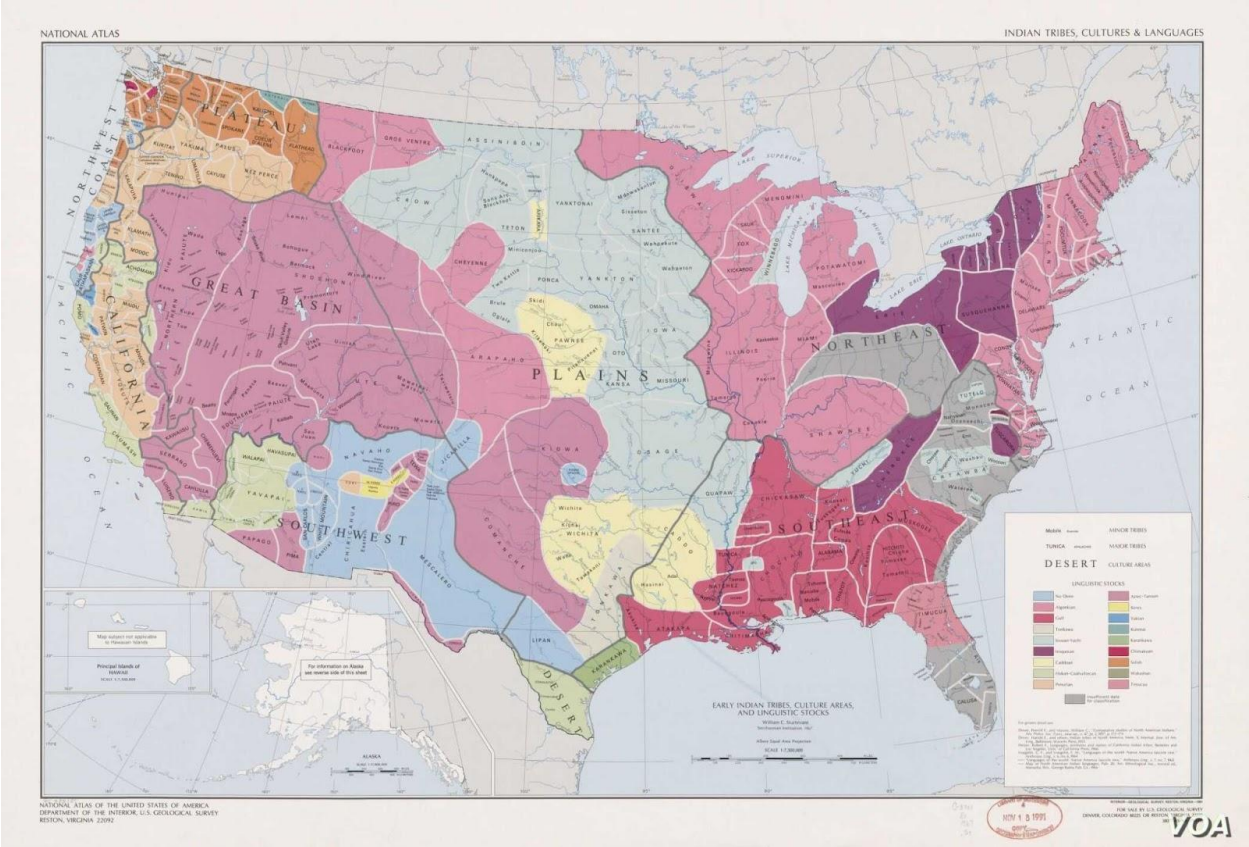


Figure: Contextual mapping of indigenous cultural regions. Plains Sign Talk was influenced by French colonization in the Pacific Northwest, as well as Canadian and English signed languages as colonization moved inland. [VOA (Voice of America), 2017]

For a long period of history, sign languages and deaf and mute individuals were not seen as valid and conflated with being inferior. This, combined with the colonization and assimilation of various indigenous groups and languages, made language acquisition, evolution, and preservation difficult for signed languages. It wasn't until 18th century France that signed languages saw more acceptance and validation within society.

Partially due to their learnings from the indigenous peoples of North America [Kelley, 2003], the French are credited with much of modern sign language and grammar. The Deaf-Mutes association in France developed and formalized sign language in the mid to late 18th century [Hendron, 2020, p 106]. This then became codified into langue des signes française (LSF) and into schools for the deaf. Prolific members of the American Deaf community later attended such

institutes and later built the foundation for American Sign Language (ASL) [Hendron, 2020]. By the 19th century, such institutions were not just in Europe, but also in the United States in the form of institutions like Gallaudet University [Hendron, 2020].

Meanwhile, other sign languages developed in colony communities all across what would become the United States in the 17th century [Lane]. By the 19th century, there were many regional signed languages based on signed languages from England and France [Lane, 2000]. Most infamously, American Sign Language, and much Deaf American cultural roots, originated from “a triangle of New England Deaf communities that flourished early in the nineteenth century: Henniker, New Hampshire; Martha’s Vineyard, Massachusetts; and Sandy River Valley, Maine”. [Lane, 2000].

Unfortunately, near the end of the 19th century the emergence of “nationalist campaigns to homogenize speed and language” emerged:

> [a] reactionary and ultimately eugenic push for ‘oralism’ swept through modernizing Western cultures, forcing educators to emphasize the use of oral speech for deaf students wherever possible. Signing, as a minority language, was discoursed as a threat to the unifying work of national culture. [Hendron, 2020, p 105]

In 1880, the Milan Conference, a major international conference on Deaf education with related authorities, declared the superiority of speech over signs, resulting in major setbacks for Deaf education and communities until the revitalization of sign language in the mid 20th century [Hendron, 2020, p 106].

This decision proliferated to both education and life for Deaf people. Instead of being taught sign language, Deaf individuals were set on paths of writing, lip-reading, and using vocal skills, or oralism. In the United States, and abroad, one of the largest proponents and influencers of oralism and its mass rollout was Alexander Graham Bell, born in 1847 to a family of elocution scholars [History Channel, 2019]. **Elocution** is a subfield of linguistics, defined as:

> “the study of formal speaking in pronunciation, grammar, style, and tone as well as the idea and practice of effective speech and its forms. It stems from the idea that while communication is symbolic, sounds are final and compelling.” [Dictionary].

This, of course, is highly audistic -- valuing spoken and sound communication over others. Bell’s work into sound and oralism was fuelled by his education but also by his mother, who gradually became deaf when he was about age 12 [History Channel, 2019]. His experiments with sound and obsession with voice and speech eventually led to his engineering of the telephone and hearing devices [History Channel, 2019].

> The speech sciences, especially as applied to deaf oral education, augured the future of modulation: control over the behavior of sound and the transmission of messages; the employment of techniques such as filtering and channeling; the enforcement of

> communication; the attachment of “defects” to sound along with the aim of eliminating them. [Mills, 2016 in Jones, et al, p 151]

However, Bell sought to further his elocution work by working towards an oralism agenda, in an effort for deaf individuals to assimilate to hearing society. This was motivated by Bell’s intense belief that 1) spoken language was superior to sign language and 2) that without integration into society, deaf people would continue to marry one another and eventually produce a deaf variety of human -- as Bell’s intention and interest in this assimilation was to eventually end deafness. While Bell never explicitly suggested that deaf individuals be sterilized or kept from marrying other deaf individuals “We cannot dictate to men and women whom they should marry and natural selection no longer influences mankind to any great extent.” [Bell, 1898], he still promoted and advanced harmful theorizing that brought Deaf people’s reproductive and personal rights into question from a standpoint of eugenics. These views are brought about in great detail in his publications “The Question of Sign-Language” and “Memoir Upon the Formation of a Deaf Variety of the Human Race”, both of which he presented in many venues.

> “It is customary among us to speak of the ‘sign-language,’ or the ‘language of signs,’ but language is that which belongs to the tongue, lingua; it is the utterance of vocal speech. In a remote, modified, accommodated sense, we may call it a language, just as we speak of the language of flowers, the language of the eyes, the language of stars, or any other non-oral method of communication.” [Bell, “The Question of Sign-Language”, 1898 in *The Educator,* retrieved from the Library of Congress]

> If then a good education, with a good command of the English language, can be obtained without any recourse to the De l'Epée language of signs, the question naturally arises, what need is there for the latter at all? [Bell, “The Question of Sign-Language”, 1898 in *The Educator,* retrieved from the Library of Congress]

> Those who believe as I do, that the production of a defective race of human beings would be a great calamity to the world, will examine carefully the causes that lead to the intermarriage of the deaf with the object of applying a remedy. [Bell, “Memoir Upon the Formation of a Deaf Variety of the Human Race”, retrieved from gallaudet.edu]

It is important to note that Bell saw the classroom as the site for this agenda, as much of oralism was done in schooling and tutoring of deaf children. Additionally, Bell actually married one of his students, a deaf woman, Mabel Hubbard [Terrell, 2012].

For the motivation of his work, contributions to assimilationist practices, and invalidating of sign language, the Deaf community views him with great negativity. Some historians argue that Bell had good intentions in that he believed “difference alone led to inequality, and he hoped that society could be improved through informed reproductive decisions” [Greenwald and Cleve, 2015]. This, of course, is still eugenics instead of seeking justice for a diverse group of human circumstances and experiences.

Intentions aside, Bell's work created sincere repercussions and linguistic injustices in the context of the Deaf cultural movement and the use of sign language. It is reported that Bell's last words, in 1922, was "No", signed to his wife -- an irony that does not go unnoticed in much analysis and criticism of his life and work [Terrell, 2012].

In the 1960s and 1970s, Deaf rights activists joined other social rights movements and formalized Deaf culture, and became recognized as a cultural and political identity in the US in 1965 [Hendron, 106]. Gallaudet University, having a resurgence of sign language in post-WW2 America as more Deaf individuals sought to formalize and further the study and use of sign language, was paramount in many of these efforts [Hendron, 106].

> "Activists made a strong case for their language, their history, and their culture as a fully normal variation on human existence; some took on *Deaf* as an identity, capitalizing it as a proper name on a par with ethnicity." [Hendron, 106]

In 1990, the ADA was passed and required interpretation, captioning, video relay services, and other communication aids for effective communications [ADA]. This is enforced at the governmental level (with all broadcasts and services having these available) but also under the ADA, individuals can request this from their employers [ADA].

Unfortunately, despite the ADA, sign language interpretation services, and other communication tools, are not always accessible for individuals who may need them. Additionally, not all interpreters and interpreting services are qualified for all scenarios. And for some Deaf individuals, closed captioning may not be suitable in all situations.

> "The condition of deafness in the world is not a personal matter only, but a social and political one; that the deaf are a marginalized identity group who deserve equality." [Hendron, 2020, p 107]

Today, the exact number of signers and Deaf people in the United States is unclear. Oralism, audism, and ableism still proliferate much of society. And yet, despite centuries of attempts to eradicate it, sign language is still very alive and diverse in today's world.

3.8.2: Linguistic traits and influences of signed languages

Signed language is a very nuanced, three-dimensional language that involves the entire body. However, most iconography and association with signed language has to do with the hands -- which is indeed the dominant way that the signs are constructed. Within signed languages, particularly ASL but several others, there are 5 parameters for each sign that are linguistic elements of the language [mtsac.edu]:

1. Handform
2. Orientation
3. Location (or place of articulation)
4. Movement
5. Non-manual expression

Perhaps most artwork surrounds the handform and movement of signed language -- this is also perhaps the most recognizable iconographically to individuals that do not sign.

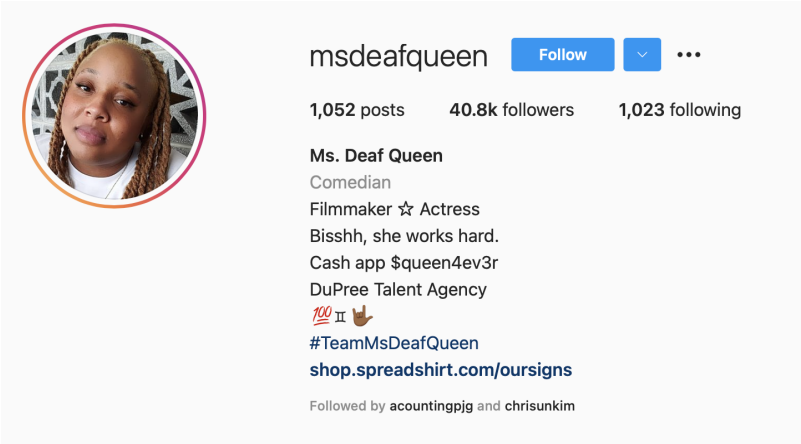


Figure: Sign language also has an iconography component embraced in modern popular culture, as seen by the sign for “love” which is an emoji that many CODAs, interpreters, and Deaf people use to signify that they use sign language

Many signs are also derived from common actions or miming of words. For example, the sign for “counting” in the logo of this project, originally came from the idea of pushing change on your hand or “counting out change”.

Figure: Logo frames for *A Counting: Sign Language.* Special thanks to Sixin Chen, Sheena Stuart-Milburn, and Ekene Ijeoma for carrying out this logo design to match the continued branding of the vocal editions.

Along with iconography and visual identity, signed languages have had written systems for centuries. The most relevant of which is Stokoe notation [Hochgesang, 2015], which has been adopted by linguists for a variety of signed languages all over the world. The Kiowa language, an indigenous language group with vocal and signed languages from what is currently the center of the United States, also has a very unique written system [Graber, 2018]. This writing system was popularized during the 19th century due to colonization to communicate with displaced loved ones [Graber, 2018]. This adaptation from signed to written language shows the linguistic complexities and importance of signed languages.

While related to vocal and written languages, signed languages often use different grammar, with the goal of brevity and directness [LSA]. American Sign Language, while influenced by English, is not used as a direct translation of English. For example -- if you were to say “I am going to the store” someone would sign “Store I go”.

It’s also important to note that signed language has a very different power dynamic and sense of ownership than spoken language. For many, signed languages are the only way that they can communicate and they often do not receive accommodations for their sign languages. As such, there is a variety of audist and ableist tendencies in written English, due to being biased from the use and development of a hearing world. It is very typical to write “Looking forward to hearing from you soon!”, an example of implicit audism. There are also more explicit idioms that are ableist and audist such as “Fallen on deaf ears”, as pointed out by Deaf columnist Sara Novic [Novic, 2021].

There are hundreds of different signed languages all around the world, and they are all unique [Parkhurst, 1998]. For example, BSL (British Sign Language) and ASL (American Sign Language) have many differences, despite both having roots in English [Parkhurst, 1998]. Due to the development of signed languages and colonization, there are many similarities within the numerical systems, however, there are some key differences that hopefully this portrait will capture as participants grow. Along with different languages as a whole, signed languages also have notions of regional dialects, colloquials, and accents [Dipasquale, 2015]. These accents may present themselves in the speed and styling of signing and even different signs for the same words or objects (e.g “soda” and “pop”) [Dipasquale, 2015].

Additionally, these accents also appear in native deaf signers and nonnative and hearing signers. This accent actually appeared in the user testing for this project, which utilized signs from Timothy Loh, a hearing person, for testing purposes. Participants doing the test were able to identify Timothy as a hearing individual signing:

> Some native deaf signers, especially those from native-signing Deaf families, can finely detect whether a signer is a native signer, an interpreter, or a post-lingual learner. Furthermore, some can sharply detect whether a signer is hearing or deaf, even if they are fluent. Hearing signers have a certain accent. [Handspeak]

3.8.3: Sign language in COVID-19

Other accessibility around signed languages has also evolved in the last year due to the COVID-19 pandemic. There is a mythos, discussed more in Section 5, that Zoom has made life infinitely easier for sign language communication. However, this again represents an ableist simplification of how sign language works.

While having meetings and telehealth appointments with closed captions has improved some accessibility, this also assumes the captions are accurate enough and that written English is a

suitable way for the individual to receive the information from the meeting. More virtual events have been occurring which also include captioning and signed interpreters in the same window instead of relying on physical seating -- this is great!

However, for many, it has increased “Zoom fatigue” [Katz, 2021] by having to look at so many places at once on a screen for long periods. One can’t look away from a meeting if that is the only way for them to receive information from the meeting. Additionally, in a video meeting, individuals are tiled, which makes some signs harder to see due to a need for seeing the detail of the hand -- and this is all assuming clear video feeds and a background that is easy to see one’s hand [Katz, 2021]. Other signs and grammar necessitate angles of the body, which can be difficult to perceive over video calls [Katz, 2021]. These limitations have resulted in signs changing and evolving to better be communicated over video -- linguistically, COVID-19 could permanently change signed languages [Katz, 2021].

For others, the use of face masks during COVID-19 has made lip reading impossible, and signs that utilize the mouth are also difficult [Callahan, 2020]. Accessible masks were developed by various organizations to help with this, particularly for interpreters [Callahan, 2020]. But in day-to-day encounters, these masks were often not present or available during the pandemic, (e.g.: many free masks are not clear or with viewing windows) adding a limitation for sign language users who were put in a situation where they made linguistic accommodations for pandemic safety measures.

3.8.4: Intersectionality around signed languages

Language, like any identity, presents itself in conjunction with one’s other identities. This thesis argues that language identities face intersection discrimination with other identities (such as race, socioeconomic status, etc).

Deaf identity and sign language are no different. It stands to reason that a White Deaf person has a different experience than a Black Deaf person. While much discourse within Deaf activism and studies seek to merge Deaf and linguistic identities around signing as cultural identity groups, intersectionality cannot be ignored:

> White Deaf people constantly remind Black Deaf people that we as Deaf people are also an oppressed group because ableist society continues to marginalize Deaf people due to their lack of ability to hear. They succeeded at shifting away from discussing how racism is being overlapped with ableism and/or audism to single issues such as ableism and/or audism. In another saying, there is no way for them to benefit from being white.
>
> ...Deaf/Hearing dichotomization without involving race isn’t logical...One of the tenets of Deaf Culture is that having a cultural Deaf identity supposedly transcends race...The huge issue with this culturally Deaf identity is based on the colorblindness ideology that excludes an account of how racist history in America has affected all of us to this day. [Player, 2020]

An ignoring of intersectionality aids in further oppression of intersectionally marginalized individuals. It can additionally erase intersectional histories within the language itself. **Intersectionality** as a framework requires holding all identities and discriminations they may face in consideration concurrently -- and thus has become a cornerstone of much modern social theory.

Black ASL, which developed due to the intersectional oppressions of Black Deaf individuals in the United States, is a linguistic artifact of intersectionality in signed languages. At Gallaudet University there is an entire line of research dedicated to Black ASL [Gallaudet.edu, BASL Project]. This language developed out of the racial segregation of Black Americans, who were historically excluded and are still underrepresented in Deaf spaces, education, scholarship, and interpretation [Gallaudet.edu, BASL Project].

Considering segregationist and assimilationist definitions, as proposed in a variety of social sciences but additionally and about this discussion within Ibhram X Kendi's antiracism framework [Kendi, 2019], the Black Deaf experience is one of painful present and past segregation. Audism as a whole is motivated by an assimilationist view of Deaf individuals assimilating to communication on hearing terms.

The severe diminishment to the point of endangerment or disappearance of indigenous sign languages in the United States is a reflection of both these systems throughout the history of the United States and its discrimination and violence towards its indigenous peoples.

Deaf individuals cannot afford to be apolitical, especially those who may not have white and cis male privilege. According to the CSD, more than 1 in 4 Deaf women will be sexually assaulted in their lifetimes [Martirosyan, 2020]. Deaf individuals are also uniquely vulnerable to violence due to the police, who do not deescalate situations of someone not responding to them, leading to escalation and violence -- these attacks motivated artist Christine Sun Kim to perform at the Super Bowl [Martirosyan, 2020].

3.8.5: Artwork and media within sign language

> ***For most of my life, I have been "spoken" or "voiced" by people on my behalf, sometimes involuntarily. As a Deaf person, voices come in many forms and are mostly functional: platforms, benefits, privileges, identities and hierarchies.***
> ***- Christine Sun Kim***

With regards to the lived experience of sign language, Christine Sun Kim is perhaps the most prolific in the context of artistic practice -- thus why so much of her practice is featured and considered in the framing of this work. In particular, Christine also bridges the gap between the power of sound and its relationship with communication and how as a Deaf person she must obsess over sound even though she uses sign language and doesn't hear it [Kim, TED].

Her art is playful and present, with drawings that interrogate the privileges that hearing people have. Kim playfully combines signs and views of them to create what she calls a "common language that all people can connect to." [Kim, TED]. She also considers the joys and nuances of sign language communication and aspects of individualism within sign language.

Other artists like Chella Man consider the intersectionality of multiple identities, in Chella's case often his queer and Deaf identity [Man]. Much of his art is about taking agency regarding communication, representation, and expression [Man] through awareness, video art, and content creation.

Scholar and artist Joseph Grigely, meanwhile, has added cornerstone thought to a variety of fields, such as body criticism, textual criticism, and relational aesthetics [Moss] [Biennial.com] and literature around Deaf artists and soundscapes. Much of his work focuses on "questions about the materialization of language and communication, and the ways conversations might be represented in the absence of speech." [FCA.org]. In his essay "Soundscapes" he discusses the physically embodied nuances of sound that a deaf or Deaf person does not partake in, since their language itself is gesture:

> We generally know what a conversation sounds like, but what does a conversation look like?
> ...
> One evening I was watching a choir on TV with my wife, Amy, and after watching for a while she turned to me and said, in sign language, "The world must look really silly without sound." [Grigely, 2016]

The scholarship of Deaf artists like Grigely has only flourished in meccas of Deaf culture and education -- which produce and inspire a variety of media and artwork around sign language. At Gallaudet University, sign is embraced not just as communication and culture but in art. Gallaudet, featured in the *Deaf U* series on Netflix, has a gallery of community artwork -- largely focused on sign language. Storytelling and movement in signing is not always recognized, instead many people focus on just direct translation, but at Gallaudet, the cultural and artistic

elements of sign are embraced. An exhibit “the Ball” at Gallaudet shows this type of storytelling in action through sign and interactive media [Gallaudet].

The Rochester Institute of Technology, home to the National Technical Institute for the Deaf, hosts a myriad of artworks and art events around sign language and Deaf culture. NTID is the largest and first technical college for the Deaf [NTID]. It has a variety of notable alumni in the arts, including Christine Sun Kim and Chuck Baird.

Chuck Baird, a painter and actor, focused much of his career on visually communicating the experience of a Deaf person [deafart.org]. His murals are extremely prolific in the context of Deaf visual culture, many incorporating sign language, including one large mural of children and adults signing together, which is now permanently featured at the Gallaudet cafeteria [Gallaudet, 2010] fitting into the aesthetic style of De’VIA, of which he was a contributor to. Baird went on to teach and engage with art in both philanthropic, educational, and professional settings throughout his life [deafart.org].

De’VIA, which stands for Deaf View/Image Art, was created in a manifesto in 1989 by 8 Deaf artists at a workshop at Gallaudet [deafart.org]

> Excerpt from the original De’VIA manifesto, 1989
>
> De'VIA represents Deaf artists and perceptions based on their Deaf experiences. It uses formal art elements with the intention of expressing innate cultural or physical Deaf experience. These experiences may include Deaf metaphors, Deaf perspectives, and Deaf insight in relationship with the environment (both the natural world and Deaf cultural environment), spiritual and everyday life.
>
> De'VIA can be identified by formal elements such as Deaf artists' possible tendency to use contrasting colors and values, intense colors, contrasting textures. It may also most often include a centralized focus, with exaggeration or emphasis on facial features, especially eyes, mouths, ears, and hands. Currently, Deaf artists tend to work in human scale with these exaggerations, and not exaggerate the space around these elements [deafart.org
>
> The name, De’VIA, evolved out of much discussion on the relative merits of an English or an ASL name. The final name, though a combination of the two, has the natural flow of ASL as the predominant consideration. [deafart.org]

Artistic and educational worlds aside, popular media and culture also have rich representations of signed language.

Throughout research and outreach of this project, it also became clear that social media influencers exist within the sign language media space. From lifestyle influencers like Cecilia Grugan, who makes content with her wife about being Deaf queer women to creative educators like Gregor Lopes -- there are a variety of influencers in this space. Both of these influencers, in

particular, partook in the user test and final 1-100 portrait for this artwork. Gregor additionally helped with the logo.

Deaf influencers have flocked to Youtube, Instagram, and the highly visual and popular video application TikTok. Additionally, interpreters, both Deaf and hearing and CODAs, have made media and art around interpretations of songs, as aforementioned in the video section of this thesis.

Additionally, the relationships of Deaf and hearing families have been central to a variety of recent media, from the film CODA [Apple TV] to an inter-hearing family on the Netflix Series Ginny and Georgia [Netflix]. Showing these modern families that are not siloed speaks towards the interconnected nature of the United States and adds representation to the real-life experiences of so many families. Behind this media are individuals like Nyle DiMarco, who was the first Deaf contestant and winner of “American’s Next Top Model” who went on to become an actor and higher-level executive producer at Netflix [IMDB].

The above is by no means an exhaustive overview of signed languages in media and art but hopes to capture some of the diversity and progress in such. From using real footage of the participants signing instead of illustrations or abstractions to building a community of recorded and transcribers, *A Counting: Sign Language* aspires to embody the organic joy and uniqueness of signing in a linguistic portrait of signed languages.

3.9: Portraits of Linguistic and Language Justice

> ***Language is personal, visceral, and powerful; it is tied to our lands, to our bodies, to our relationships, and to our knowledge. Every time we speak or sign in our particular accents and dialects, syntax and rhythms, cadences and inflections, we identify ourselves and bring social history and personal experience with us. When we come together to dialogue, it is important that we are able to express ourselves in the language that most fully conveys the depth and nuance of our hopes and ideas, our frustrations and questions. [Antena, 2012]***

A Counting seeks to create a linguistic portrait of what a truly united society would look and sound like. Such a society would be linguistically just, but there are important distinctions and elaborations regarding the various discriminatory practices that prevent language justice in today’s society, as well as the differences and relations between linguistic and language justice.

3.9.1: Audism, Linguicism, and Ableism

Previous sections have expanded on audism and ableism and how they affect sign language and its community, particularly Deaf and hard of hearing individuals. This linguistic discrimination against sign languages, which is part of audism, additionally connects with **linguicism.** This term was originally coined by Tove Skutnabb-Kangas in 1988 to call attention to the denigration of minority languages [Bauman, et al, 2013]. Often educational institutions are the first site of this type of discrimination [Bauman, et al, 2013].

In a briefing by H-Dirksen Bauman, Scott Simser, and Gael Hannan to the Canadian Hearing Society, it is presented that:

> The distinction between medical, social, and cultural models is key to understanding the following key terms, ableism, audism and linguicism. Audism may be seen as both a subset of ableism and of linguicism, thus forming its own unique dynamic of discrimination. [Bauman, et al, 2013]

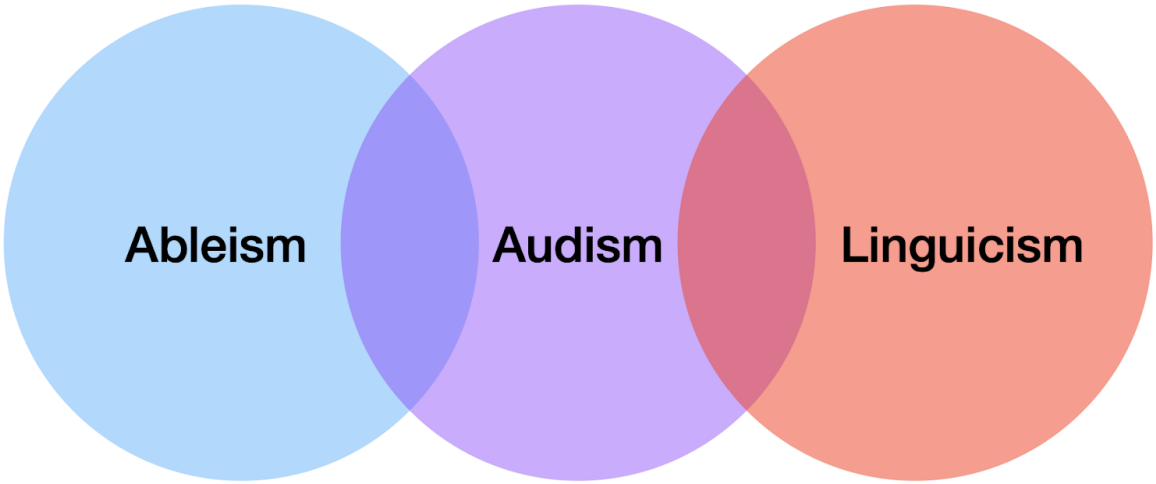


Figure: Adapted from the report "Beyond Ableism and Audism: Achieving Human Rights for Deaf and Hard of Hearing Citizens" showing the interlinking relationships of ableism, audism, and linguicism.

Considering these networks of discriminatory systems, it stands to reason that there are nuanced justices to combat them, particularly in the form of linguistic and language justice.

3.9.2: Linguistic and Language Justice

But what is linguistic justice and language justice? How do they differ and how can they inform linguistic portraits?

For this thesis, a **linguistic portrait** is any multimedia artifact centered around language(s) that continuously conveys a contextualized message through the speaker(s) of said language(s). In the case of this project, in particular, *A Counting: Sign Language* is a **gestural portrait** -- an artifact that is meant to do this for signed languages. This distinction is particularly important in that many linguistic frameworks do not acknowledge signed languages.

Both terminologies around this artwork differ vastly from **language portraits,** which are educational tools used within the linguistics field and are not inclusive of signed languages [Busch, 2018]. This is an example of linguistic and language education's exclusivity at large.

Arguably, the social practice seeks to materialize artistic and public practices in an effort towards justice. Enabling, creating, and exhibiting an inclusive **linguistic portrait** is how *A Counting* engages with linguistic and language justice, which are also distinct in goals and framings.

Linguistic justice is canonically used in the literature regarding Black Language and Black Liberation, as cited by linguistic scholar April Baker-Bell, who coined the term, in her research:

> It [linguistic justice] is an antiracist approach to language and literacy education….dismantling Anti-black Linguistic Racism and white linguistic hegemony and supremacy in classrooms and in the world. [Bell, 2020, pg 7]

Baker-Bell explores eradicationist language pedagogies, which seek to erase language and language identity from marginalized students, particularly Black students in the American public education system. She outlines the following approaches towards language pedagogy, which could be applied to any language identity:

Table: Baker-Bell pedagogies summaries and examples

Pedagogy	Summary	Examples
Eradicationist Language Pedagogies	These policies eradicate marginalized languages by setting them up as inferior and replace them with mainstream White English.	Residential schools Audism within Deaf communities
Respectability Language Pedagogies	This approach respects the marginalized language and acknowledges that it “should be validated, affirmed, and respected”. In this approach though, the goal is to use a marginalized language as a “bridge to learn White Mainstream English” [Bell, 2020, p 28]. This approach “adheres to politics of respectability” [Bell, 2020, p 28].	English Second Language education Exact signed English
Antiracist Black Language Pedagogy	“Black Language is placed at the center of this approach to critically interrogate white linguistic hegemony and Anti-Black Linguistic Racism. The end goal of this approach is to dismantle Anti-Black Linguistic Racism and students' internalization of it.” [Bell, 2020, p 28]	As a newer framework

This thesis is by no means centered in education or the classroom, it is important to understand the world of linguistic justice in the context of how these languages became marginalized and oppressed. This is seen with indigenous languages through abusive residential schools, which served as a site of eradicationist education of various indigenous languages [Lemay, 2021]. Additionally, linguistic justice can be seen as a method to work against the aforementioned definition of linguicism.

Language cannot proliferate without education -- both in and outside of institutions. Signed languages in particular have faced both eradicationist and respectability language pedagogies and these have severe impacts on users of signed languages on individual and systemic levels.

From Deaf individuals who have families that will not sign with them, to second or third-generation immigrants who lose their ancestral language. Part of the conceptual grounding of counting for this work is that numbers are one of the first things one learns when learning a language -- and this cannot be considered without acknowledging the role of education and abuse, eradication, or inability to access such education around one's language.

This wide view of intersectional **linguistic justice** regarding languages that have been historically oppressed, marginalized, and stolen over time from a variety of groups is important for all lines of work for *A Counting.*

Therefore, for the context of this thesis, **linguistic justice** is more widely defined as:

> An intersectional system for approaching the treatment, acquisition, and preservation of marginalized languages in society, such that one can gain access to acquire and utilize their language with dignity and respect.

Linguistic justice is fighting against linguistic oppression — in particular with signed languages in which many Deaf and HOH individuals faced eradicationist measures (told not to sign, not taught, and audism seeking to prioritize hearing and speaking over signing) and language deprivation, another method of linguistic oppression.

The artistic group *Antena* has theorized extensively on **language justice** as a social justice issue and on artistic interventions to achieve a more language-just society. To envision a language just society is to envision **multilingual spaces.**

> When we say multilingual space, we don't just mean a space where many languages are spoken, but spaces where there is a commitment to equality among languages, as well as a resistance to the dominance of any one language in the room. [Antena, 2012]

Unlike linguistic justice, **language justice** refers to the right of someone to *communicate* in a space, particularly in the language that they feel the most comfortable in [Antena, 2012]. This can be accomplished through multilingualism in societies but also through a variety of interventions to current communications, ranging from broader access and visibility to interpretation to utilizing technology to enable more streamlined translation tools [Antena, 2012]. And while these concrete interventions are important, **language justice** also needs to be accomplished by understanding the importance and intersectionality of language and how it can empower or disempower someone, and value all languages, including signed and rare languages, which have been historically undervalued.

Therefore, for the context of this document, **language justice** is defined as:

> A justice system that values and enables all languages, allowing everyone to communicate in the language that they are most comfortable in free of discrimination as they navigate society.

Linguistic justice seeks to dismantle the oppressive systems that make **language justice** difficult to attain in the United States, which would enable many intersectionally marginalized people, such as signers, to communicate and live their lives in safety and dignity.

3.10: Relationals

> **Almost every work of serious contemporary art recapitulates, on some explicit or implicit level, the historical sequence of objects to which it belongs...Yet the organized discipline of the history of art remains largely blind to the products of this interest and entirely sheltered from the lessons that might accrue from them. [Thomas Crow, 1995, from Buchmann p 57]**

The previous sections position *A Counting: Sign Language* within different artistic practices and practitioners and add background and new discourse to the variety of concepts, cultures, histories, and frameworks that this work ruminates.

As apparent, much of this is extraordinarily intersectional and multifaceted -- an entire hyperspace of possible framings and considerations. This of course is ironic in that much of this work, including future participation, is open.

This thesis does not hope to limit or exhaust the discussion, both of this work and others mentioned, but rather to help frame this particular work and to propose new framings regarding language and linguistic justices in conceptual, participatory artworks, particularly for signed languages. The hope is that much of this could be imagined in search of framing artworks around other justices and identities.

The last theoretical proposal of this thesis is a higher level one regarding its connections to various histories discussed in previous sections. Particularly how social fields and understandings (particularly art history and 20th-century American history) engage with history and experience.

It has been established that languages are part of one's intersectional identities and individuals can be systematically oppressed through linguistic injustices. This can be viewed both in the present, on micro and macro levels, and in historical fashions. Furthermore, opinions, theories, and artworks around this exist across this temporal axis.

Much of this is directly with respect to how history is taught and understood at a common level, especially in the United States. Across the US, only 87% of schools mandate education regarding indigenous peoples, and only mandated in pre-1900 context [Whisnant, 2019] and where the teaching of the history of systemic racism is a large debate in 2021. There was a

recent TikTok conspiracy trying to rewrite the history of Helen Keller, a deaf-blind woman, promoting the ableist ideals that she could never have accomplished what she did during her life and spreading false information about the nuances of disability experience [Onion, 2021]. These pipelines of conspiracy threaten even extremely recent history, as seen in early 2021 when misinformation regarding the 2020 election led to an insurrection on the Capitol.

Much in the same way that linguistic pedagogies can be eradictionist and enable further oppression, so can any history.

This thesis would like to postulate that there is a mythology that we live in a post-historical society. Simply by the definition of post-historical, we do not. History is continuing and is a direct network effect with previous histories, now at a new scale in the 21st century.

However, the belief in post-history gives power to mainstream groups to not engage with past atrocities and their current ramifications.

The counterargument here is that history is relational. Relational in that, depending on one's position in society, history is understood and enacted differently in one's life. And different common historical understandings can become popularized due to their relationships with powerful identities in societies and with cultural changes.

Histories are relative in the frame of their documentations and biases, but furthermore can be tools of empowerment and oppression. Similar to the definition of aesthetics posed within disability aesthetics relies on the juxtaposition of concrete truths between multiple parties -- what response does a presented history cause someone?

It is important to note that while Joe Blosser wrote on relational history regarding economics and Adam Smith's nonlinear economics histories [Blosser, 2020], for the focus on this document relational history, much like relational artwork, is the idea of framing theory around human relations and social contexts -- in this case particularly the complex system of language and quantification within the power dynamics of intersectional identities.

Consider the Census, which is a historical artifact, in a relational view. One could consider the relational structure of racial categories and how this may affect individuals and communities, while perpetuating already strained relationships between them and the government administering the Census. One could further consider the distribution of the Census and how this can affect legislation and resources as bad actors can weaponize it to expand the gap between the lower and upper classes even more.

Relational history and relational artwork pose a unique framing for this thesis -- one that intersects with power but offers a framework to work towards agency and dignity within artworks and representations of different identities. While *A Counting* is not a relational artwork, it is part of the relational history of language and quantification of intersectionally marginalized individuals in the United States. Even seeing that the first edition did not support signed

languages is evidence that the work is not even post-historical in itself -- after all, there are fully written and tactile language systems that the artwork is not currently compatible with.

This framework of relational history also recognizes that as cultures shift, so can historical narratives and they can be rewritten by members of those identities -- such as how Christine Sun Kim has rewritten and expanded on the relationship of sign and sound art and how it intersects with the history of her life growing up during a time of ASL expansion in the United States.

While language is relational to history, it is more than symbolic in this work or other artworks. Rather language is a construct, yes, but also an identity and system in society that interweaves with other components of someone's resources, identities, and methods of interacting with the world. This is not to eliminate, however, the experience one feels when seeing signed language in these video artifacts. Language as a system and identity within history is important, but language as an experience in these artworks is equally important.

Language is a complex system parametrized by linguistic rules and semantics -- it is a model of communication, which is experienced by all and experienced differently in this artwork -- as not just a count, but a meditation of language unity and openness to signed languages, which were originally not part of this count. This work *models* the fundamental idea of language and its considerations and consequences in modern society.

> How, and how differently, do art and science model experiences? [Jones, et al, 2016, p 13]

This is the question from Caroline Jones' 2016 essay *Modelling.* The argument Jones makes in this essay introduces the book "Experience" which focuses on artwork in the context of cultural and mass experience and thought around artifacts and the meanings imposed on them (thus the modeling points from Jones' component).

> Modeling is a practice common to both artists and scientists, and a customary activity of humans in general; our modeling behaviors distinguish us as a species. We model ideas, we model clay, we model with pictures, texts, clothing. [Jones, et al, 2016, p 24]

The work at the center of Jones' essay on modeling is a geometric film work towards causal attribution in social perception -- seeing how individuals will project story and emotion towards inanimate shapes on a film.

> Experience is our heuristic...As art historians, we are particularly concerned to make space for art in these discussions -- art as culturally produced modality for instigating special kidneys of experience that both stretch cognition and model it. [Jones, et al, 2016, p 15]

And this is where the subjectivity, perception, and projection intersect with this model and this artwork. And where, like history, depending on one's experience, a different experience emerges from an artwork, albeit steered by a common experience and precedents that lead to that artwork.

> Art perpetually shifts in culture because what art models necessarily changes over time, and what we want art to reorganize us into, also changes. [Jones et al, 2016, p 26]

It is appropriate to consider this framing with respect to the sensory experiences of language -- of sound and reading in the voice portrait edition, and of seeing and understanding signs in the video portrait edition. To consider how the histories of signed languages and relevant artworks shape the experience of this model of language as a representation towards a more just and united society. Or, as Jones says more concisely: "Art is the kind of modeling that can effect the change." [Jones et al, 2016, p 33].

And thus, this thesis introduces the artifact *A Counting: Sign Language,* a public, participatory artwork and linguistic, particularly signed, portrait of the United States.

4. Technical Contributions

This chapter will outline the objective technical contributions of this thesis as a whole and their novelty and merits within current solutions. This section covers the technical contributions of this thesis which includes:

1. Cross platform recording interface
2. Cross platform transcription interface
3. Web application to host and integrate all these components
4. Process to scale, crop, and center raw video footage for a final, consistent video portrait
5. Methodology for automatic whtie color balancing for a variety of skin tones and video conditions and settings
6. Assembly of 100 unique videos in a final video portrait

More results, analysis, and room for improvement regarding these solutions are in Section 6.

The entire thesis software system is diagrammed in this higher level diagram below:

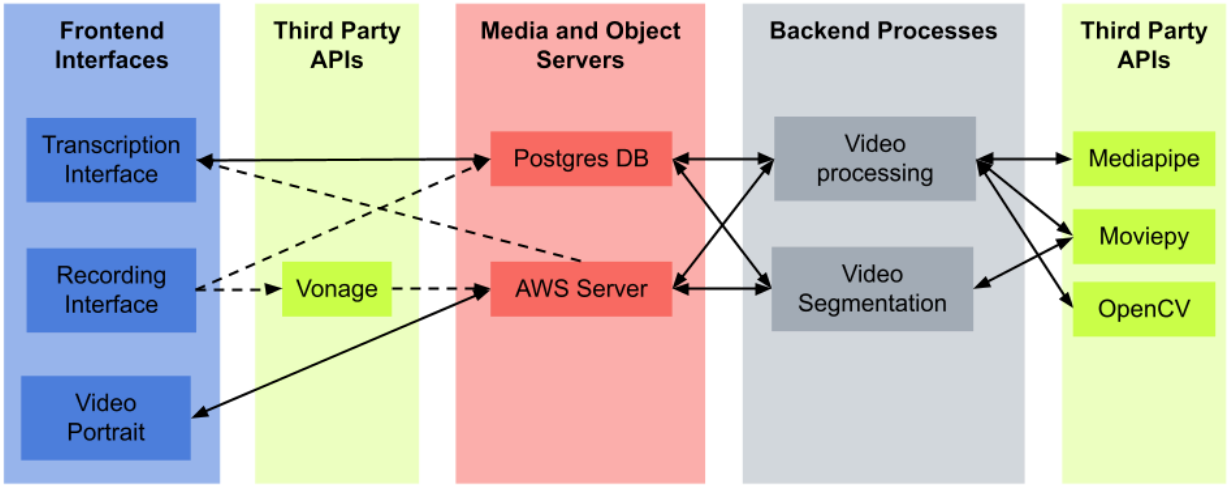


Figure: Higher level system diagram of the A Counting: Sign Language system. Here dashed lines are one way information flows and solid lines represent two way information and logic flows in the implementation.

4.1 Recording interface and software

The first necessity of A Counting: Sign Language was to develop a recording interface and system for participants to record themselves signing from 1-100 on any device on a variety of network speeds. These videos then had to go to our servers for processing in a secure fashion.

4.1.1: Challenges in Cross Platform Video Recording

Overall, streaming and recording video via the web across platforms (desktop, mobile, tablet) at a low cost is a non-negligible task. The video must be rendered in real time on the web page as

a canvas object, as well as seamlessly routed to a media server. There's a reason that companies like Twitch are so successful, as this is what they specialize in. This thesis utilizes a low cost video conference service API within a Python web application to achieve streaming of the recordings to the database.

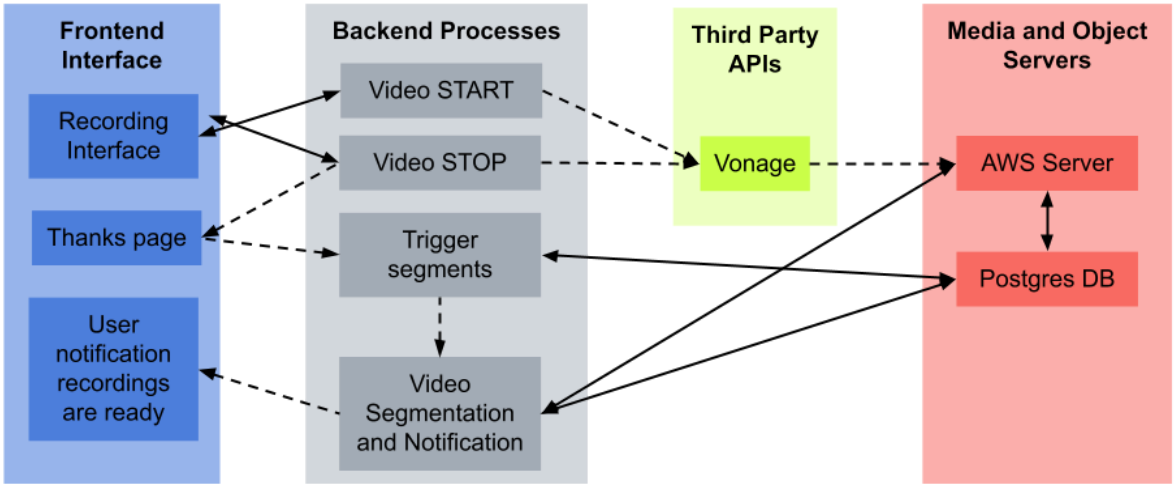


Figure: Showing the process by which the A Counting Sign Language server initiates an API call via Vonage to start streaming and recording video on the website to be saved in real time in the AWS server.

At the most basic level, this process requires a client and server connection. A stream must be created so the user can see their video feed and a recording must be initiated so that this stream can be saved to a server for future access.

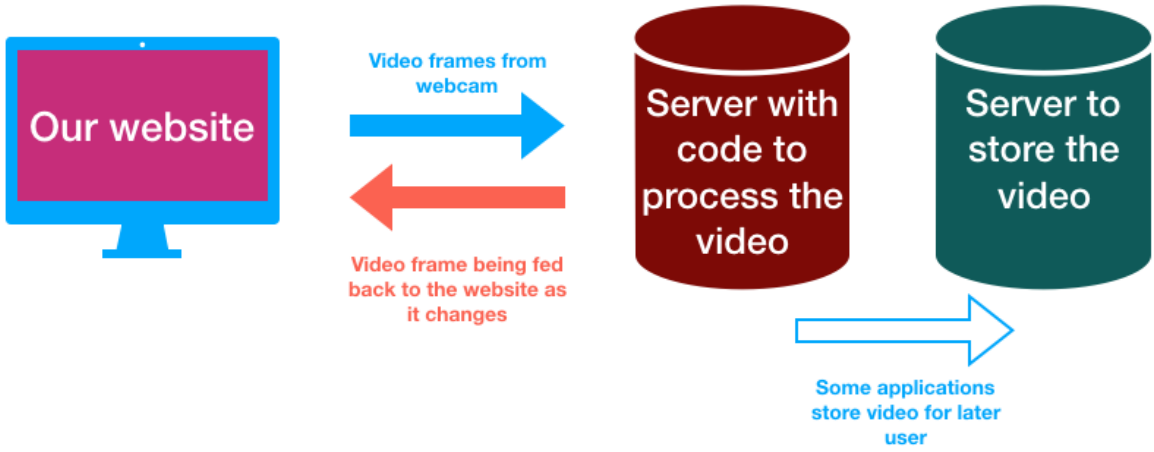


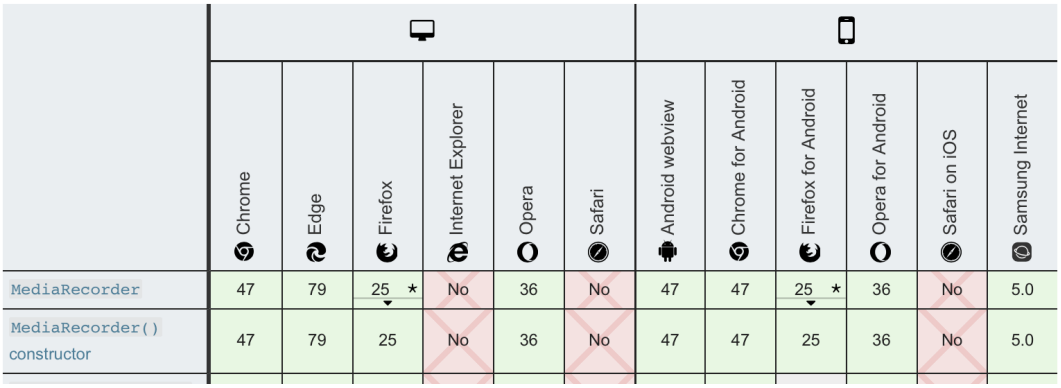
Figure: Figure showing a standard pipeline for video streaming on a website.

This can be done either on the client or server side -- on the client side, videos are recorded locally and then processed at the end through an upload. This is both more memory and time consuming and not suitable for this project as the goal is for users to be able to participate on any device and in slower networks. Server-side recording routes the stream directly to the server, taking the processing and memory away from the recording device. This additionally

adds security, as the video object, and its corresponding metadata, isn't being sent at the end across the network. Thus, more appropriate for this project.

Since server-side architecture solves any memory and efficiency access issues, the largest access issue is browser and device compatibility. Many of these complications are largely due to internet implementation around how video is streamed and recorded on the web. Javascript and HTML do not have native functions for doing this, so third party APIs have grown in various development and browser ecosystems. In particular, WebRTC and RecordRTC (which implements WebRTC) are exceedingly popular [RecordRTC.org] and rely on MediaStream and MediaRecorder to serve and store video, respectively [Mozilla].

However, there are issues that arise from depending on WebRTC and MediaRecorder [WebRTC][Mozilla], which were developed primarily in Google Chrome and Mozilla infrastructures [Mozilla]. These infrastructures aren't supported on all browsers and all mobile devices, particularly Safari on Desktop and iOS, as seen in compatibility documentation below. This creates tension when designing an infrastructure for all devices, which was exceedingly important for this project.



	Chrome	Edge	Firefox	Internet Explorer	Opera	Safari	Android webview	Chrome for Android	Firefox for Android	Opera for Android	Safari on iOS	Samsung Internet
MediaRecorder	47	79	25 *	No	36	No	47	47	25 *	36	No	5.0
MediaRecorder() constructor	47	79	25	No	36	No	47	47	25	36	No	5.0

Figure: Compatibility chart for MediaRecorder API from Mozilla, which is the most common architecture for recording video on the web.

One alternative is to exploit HTML5's canvas by treating a video stream like individual still camera frames into an animation. This is not realistic, though, for real time camera footage and can drop frames and footage in renders, especially in slower connections [StackOverflow, 2013].

There are a variety of expensive third party services available, such as CamTag in Heroku, but these are extraordinarily expensive per month for just a few hours of HD video a month. Lower cost "hacks" and newer libraries [VideoJS] are still often not stable across all platforms as well or advertise compatibility that require the device to be set in "development" mode, which is just unrealistic for most user experiences.

Building a solution from scratch that would resolve all of the above tensions was out of the scope of this thesis, and many options were tried, resulting in one low cost solution of the third party API Vonage.

4.1.2: Solution to Cross-Platform Video Recording

After many experiments, the solution determined for this thesis was to use the Vonage Video API [Vonage] for Python and JavaScript and integrate it within the existing Python and AWS ecosystem for streaming and storing video files.

The Vonage Video API was formerly OpenTok from TokBox, which was acquired by Vonage. It was primarily used for health applications and customer video chat support. While not a free service, it is low cost at $10/month for 2,000 minutes of HD video [Vonage]. This was reasonable within the research budget of the POetic Justice Group and as a justification and tradeoff for a low-cost system of this type, especially considering labor trade-offs and security advantages.

For this thesis, the archival structure for web therapy apps, a common Vonage use case, was adopted for the recording interface [OpenTok] and subsequently connected with the project's AWS storage within the Vonage video account.

The endpoints to start and stop archiving are hosted within the A Counting: Sign Language DJango application within Heroku. They are triggered within the Javascript code for the recording interface.

For the frontend, Vonage treats the video as an HTML DOM element similar to a canvas. The API has its own default UI, which can be easily overwritten using the methods outlined in its documentation [TokBox]. These parameters can be set in the JavaScript to record just video or just audio and the size of the video's appearance. However, Vonage will store the device's default camera video based off the resolution given to its video object -- these size constraints for the video object are really for what the user sees.

This structure allows for reliable cross-browser and cross-platform recording at a low cost with the security and stability of large third-party software services. Furthermore, it allows for customizable UI overlays and responsiveness across screen sizes, as seen in section 4.1.4.

4.1.3: Comparison of other solutions tried

Many solutions were tried in coming up with this final pipeline, their issues are outlined in the table below:

Solution	Summary	Issue
CameraTag	Third party professional software works amazingly and in the Heroku development ecosystem.	Far too expensive.
Google getUserStream	Google's video API works across a variety	Not cross browser still - WebRTC base, which does not work on Safari mobile correctly, as it wants the user to record and upload instead of streaming in the browser.
HTML5 canvas frames	Saving individual frames and looping them back to the HTML canvas	Far too slow and time consuming to implement for these purposes.
MediaRecorder	An API from Mozilla to record and stream video media on a variety of platforms and use cases.	Not cross browser (see screenshot above)
WebRTC	An open source project from Google for real time communication capabilities	Not cross browser for streaming playback -- especially for Safari
Videojs-record	A Javascript plugin for recording and playing audio and video files on the web	Videojs-record works on all browsers for playback, but recording depends on the MediaRecorder API from Mozilla, and thus is not cross browser compatible

4.1.4: Recording interface design

To understand the design process and the interaction flow for the recording interface, it is important to understand the user experience of the phone hotline edition of A Counting. In the phone edition, users call and count in increments of 25 with prompts between them, then leave a comment after these 4 segments.

To mimic the phone line, the system has the user sign in increments of 25 (1-25, 26-50, etc) and pause at each sign for the purposes of consistency in the final portrait and for segmenting between individual signs. Each individual sign is timed for 3 seconds.

Users would then tap or click the screen when they were done with the increment of 25 and trap or click to start the next one. This allows users to pause and not have to sign 1-100 in one go, which could be tiring. At the end, like the phone edition, they may leave a comment.

These are all done in instruction overlays that are simple HTML DOM elements (text and CSS box with border) that are hidden and shown due to a JavaScript timer in the recording code. The z-index of the CSS of each element is set to make them appear above the video in absolute positions. Using the Bootstrap CSS framework, these positions stay absolute across screen sizes so that the interface looks uniform across screen sizes, as seen below.

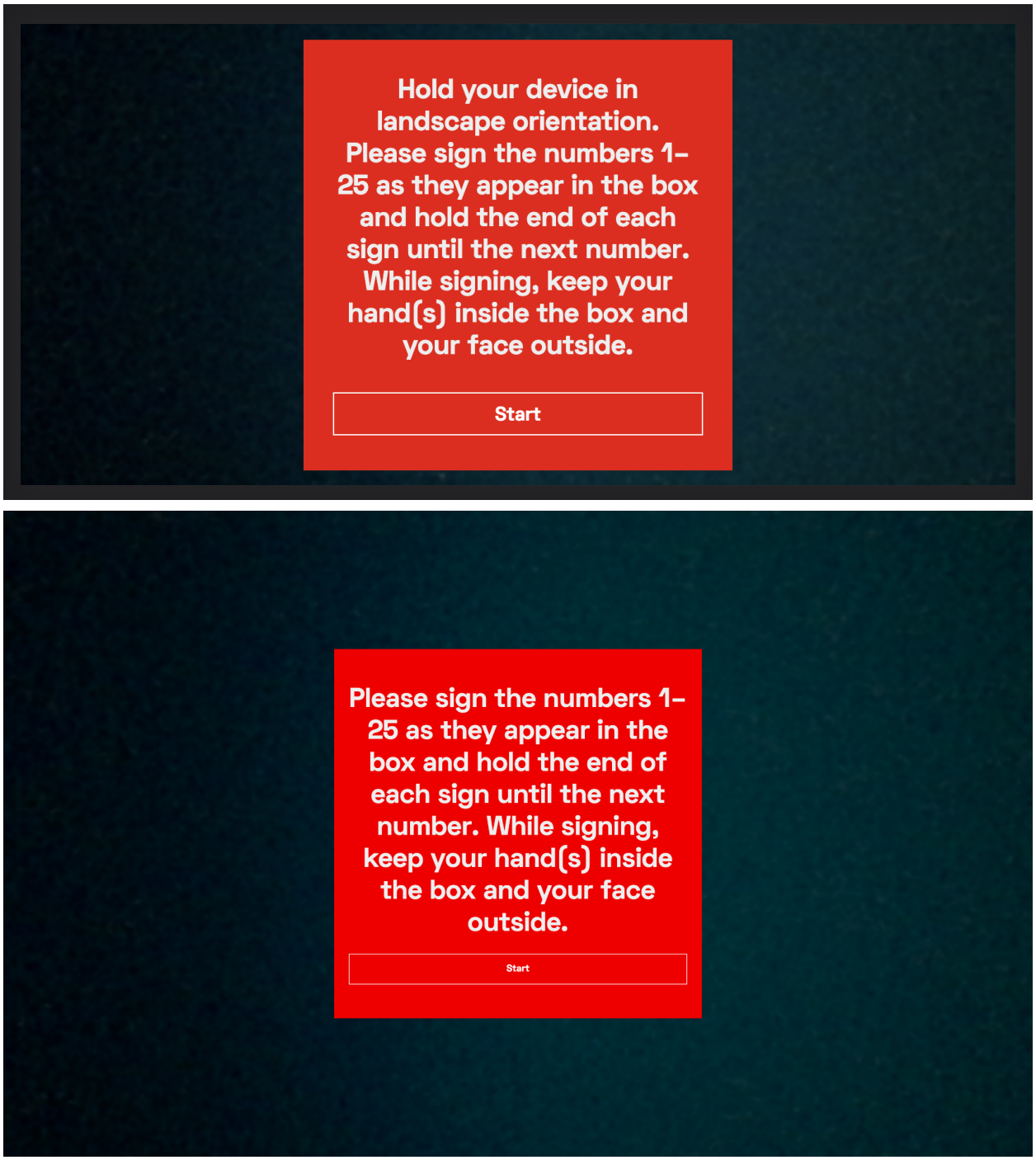


Figure: The first prompt of the recording interface in both mobile (top) and desktop (buttom) screen sizes.

A user test, described in Section 5, was conducted to determine pausing as an in-between state between individual signs as well as the 3 second timing, which was also informed from recordings of ASL numbers.

Originally, when developing the storyboard for the final interface and flow described above, it was considered to have signs go in pairs, but this would be far too much clicking.

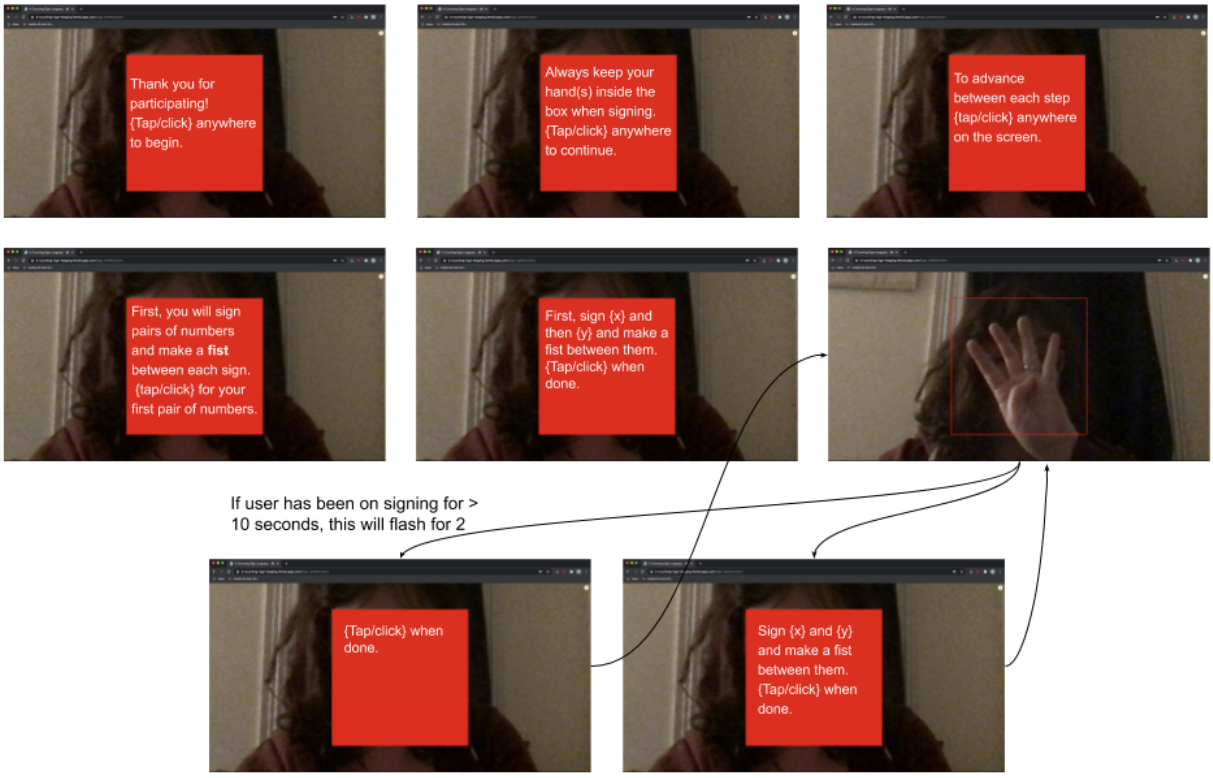


Figure: A preliminary storyboard from the designing process in fall 2020

Thus, a simple timed number prompt for each number was introduced in the box for the fall storyboards, with text boxes only for occasional instruction.

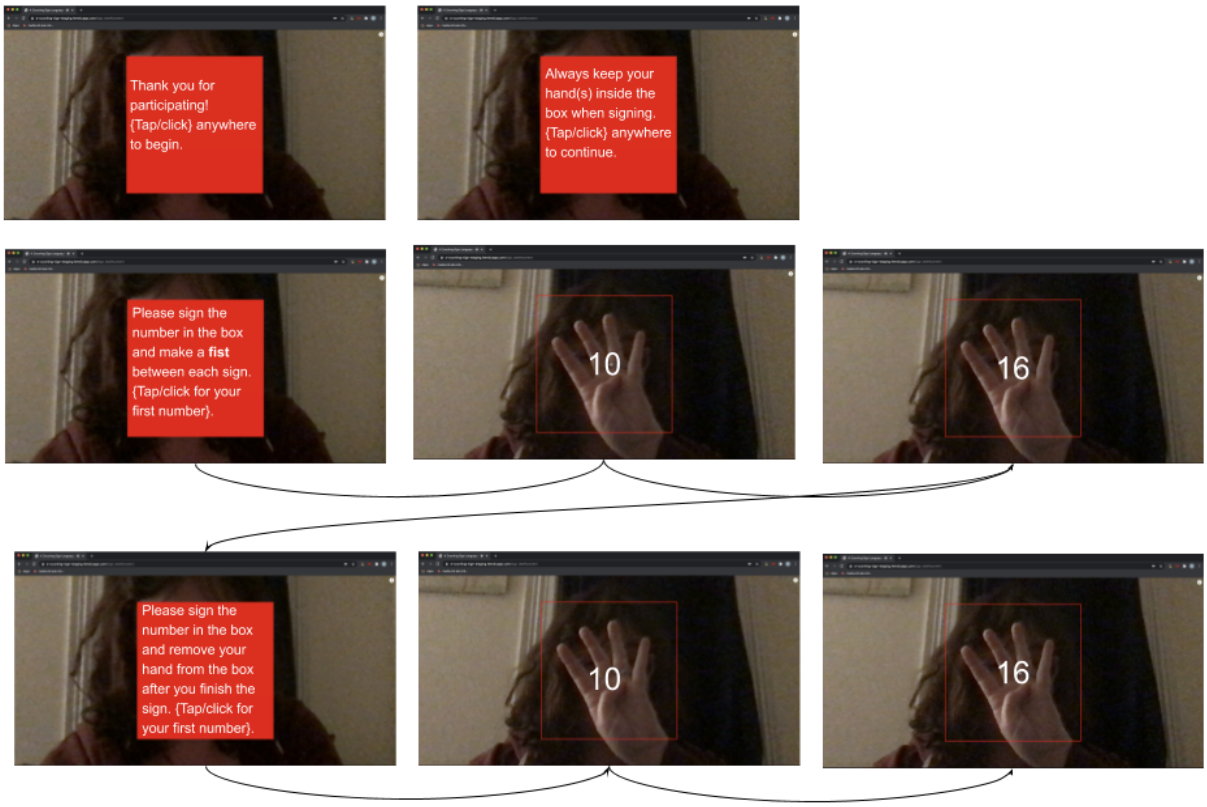


Figure: A preliminary storyboard from the designing process in fall 2020

This, of course, was modified with copy and sequencing of interactions for the final flow above.

4.1.5: Information pipeline and security

In order to ensure the security of the video files themselves as they travel from the recording interface's stream to our server and for other processes in our pipeline, video files are stored directly from Vonage archiving into our AWS files.

These AWS files are never transferred directly through POST requests, but only hash identifiers that may refer to the count objects that they refer to.

4.2 Transcription interface and software

The transcription schema for A Counting: Signed was in many ways easier than the schema for A Counting: Spoken because there are no written transcriptions for the signs themselves -- only numerical values. Users watch the video and select the numerical value or if the sign was invalid; these are easily tracked in a database.

There is an additional video of feedback at the end of the counts that some users chose to leave -- users are meant to transcribe these in written English.

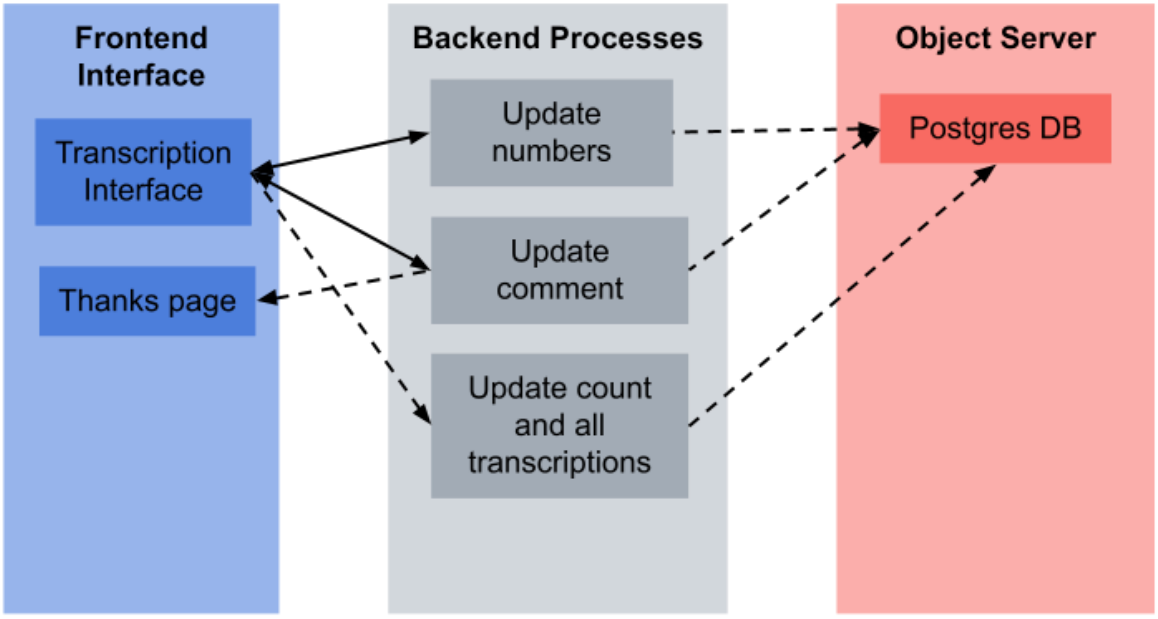


Figure: System diagram of the transcription interface and processes

4.2.1: Multiple video embeds and loading

The main technical hurdle involved with the transcription system is the fact that instead of 100 audio files, every count has 101 video files (1-100 numbers and an additional allotted for any comments). Loading multiple video files, and their poster frames, on a browser page isn't easy to do, especially on slower networks and for mobile devices. Furthermore, each video file needed to be loaded within the same DOM element so that they could be displayed in the interface grid as seen in 4.2.2 below.

This is primarily because embedded videos still require the page to make additional requests for files and resources needed for video playback, even if all of the visible videos do not need to be played at the same time and even if the files themselves are not very large. This can slow down the entire page significantly.

Most popular video platforms load thumbnail or GIF images and then when the user clicks one sends them to a separate webpage with a video playing. But this solution would not be

appropriate for our purposes, as the goal is for users to play multiple numbers in the same transcription interface -- not to go to different pages.

Furthermore, using Youtube's API, which allows developers to embed multiple Youtube videos on a single page rather seamlessly wouldn't work either as listing the videos on Youtube wouldn't fit well into our pipeline, since YouTube takes longer to render and process any uploaded video to their platform.

Overall, two main goals had to be accomplished in order for the transcription page to load in a fast and accurate manner: they needed to be loaded in a controlled manner and the files would have to be compressed.

The first method to prevent the page loading to speed up is to not load the video files at all until the document itself is loaded. This is why there are loading patterns on the web where much of the page appears before multimedia content is ready to play or be interacted with. A single loader was added to the interface in order to enable this loading pattern.

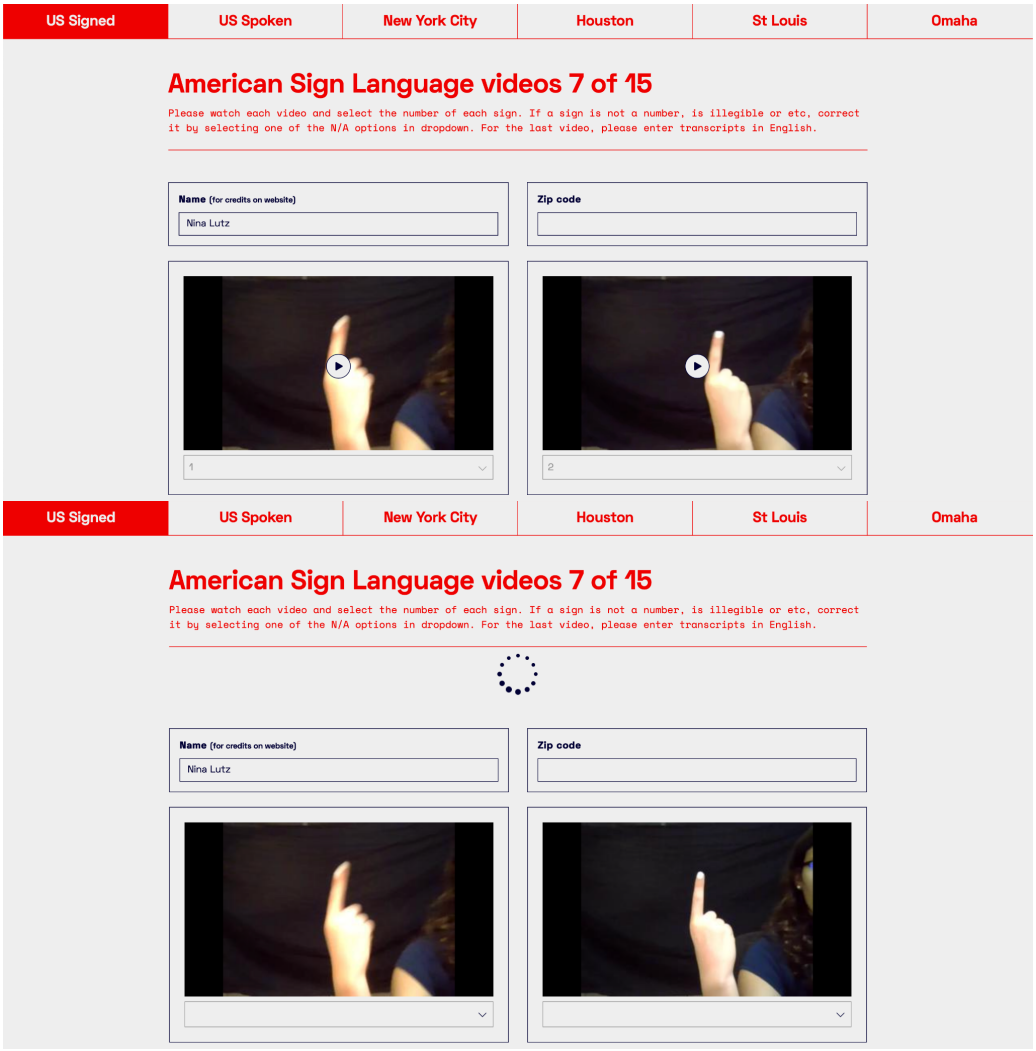


Figure: Loader only disappears when the video file is loaded and ready to play, until then the poster is visible.

Along with when to load the videos, the next question was how many videos to load at once. This had a lot to do with the video's file size (and therefore load time) [W3.org], expressed in bit-rate and file sizes. Bit-rate is the amount of data the video uses per second when it's played back [W3.org]. Bit-rate is set during the exporting process and, in most cases, a higher bit-rate will result in higher quality video, but the cost is a larger file size [W3.org]. This was desired for the portrait, but not needed for transcription purposes. In a grid, the videos are much smaller. Therefore each video file has 2 versions: an HD raw footage and a compressed video file for transcription purposes.

Compressing videos for transcription was done in moviePy, the library used for video processing in this project. After experimentation, a good size for the video files for the transcription interface was 30% of the original size of the HD raw footage from Vonage.

However, even with compression and a loading method, the files were still too large to have all of the videos loaded at once on the page in a timely manner.

Fortunately, it is not necessary to load all of the videos at once, as all 101 would not fit on the same page visually. Thus, videos are loaded in chunks. At the start of the interface, *n* videos are loaded. After *n/2* videos are played, the next *n/2* are loaded into the page. In the case of the current system, this *n* is 10, which is still too much for one viewpoint, but this gives a margin of error for any network slow downs. Similar to how many social sites structure their newsfeed load methods.

It is worth noting that that interface itself is still requesting all of the files in the background, since all 101 are declared in the HTML document of the page.

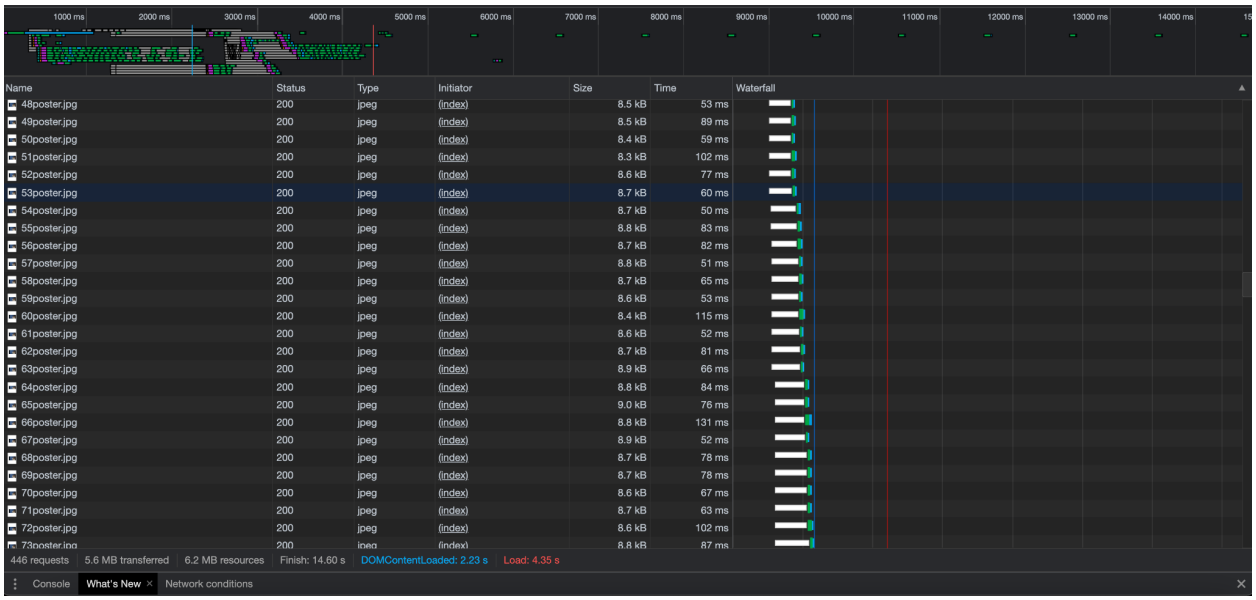


Figure: Waterfall of multiple downloads in the background on a transcription page as seen in Google Chrome's network inspector.

While this chunking method worked on desktop browsers, a variety of mobile browsers restrict the number of video elements allowed on a DOM document. This is for cellular data and security reasons. A better solution for this is to have pagination for mobile browsing experiences. This ensures that each video can be transcribed and that each page will be able to load and not freeze or crash.

4.2.2: Interface design

The interface design for the sign language transcription interface is the same design as the vocal edition, the components and layout meant to mimic each other.

Figure: Desktop interface for spoken (top) and sign edition (bottom), showing the extensive visual consistency of colors, fonts, and UI elements such as dropdowns, play buttons, and text fields.

The main differences, of course, are that the signed edition has videos instead of audio files, and thus takes up more layout space per row. Because of this, a progress bar was added to help users understand where they are in the transcription process. And after being played, videos are grayscaled.

Due to restrictions regarding having multiple videos on the same HTML page, as elaborated in the previous section, the transcription interface on mobile actually uses pagination with the same button styles. This is another good use case regarding the progress bar.

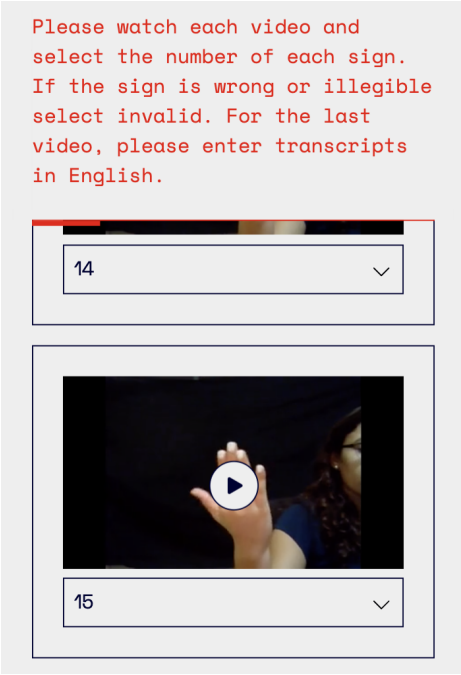


Figure: Phone version of transcription interface, featuring the progress bar

4.2.3: Information pipeline

As with the recording interface, there are never any video files that are sent via POST requests or throughout the information pipeline of this system. Instead, when a user transcribes numbers in the interface, a POST request is made with the hash ID of the count, as seen in earlier diagrams. This prevents sensitive information from the videos or the database being leaked while users are utilizing the interface.

4.3 System Architecture

Like the A Counting vocal edition and the Freedom Radio project within the Poetic Justice ecosystem, A Counting: Sign Language uses a DJango Python web application and Postgres database deployed through Heroku.

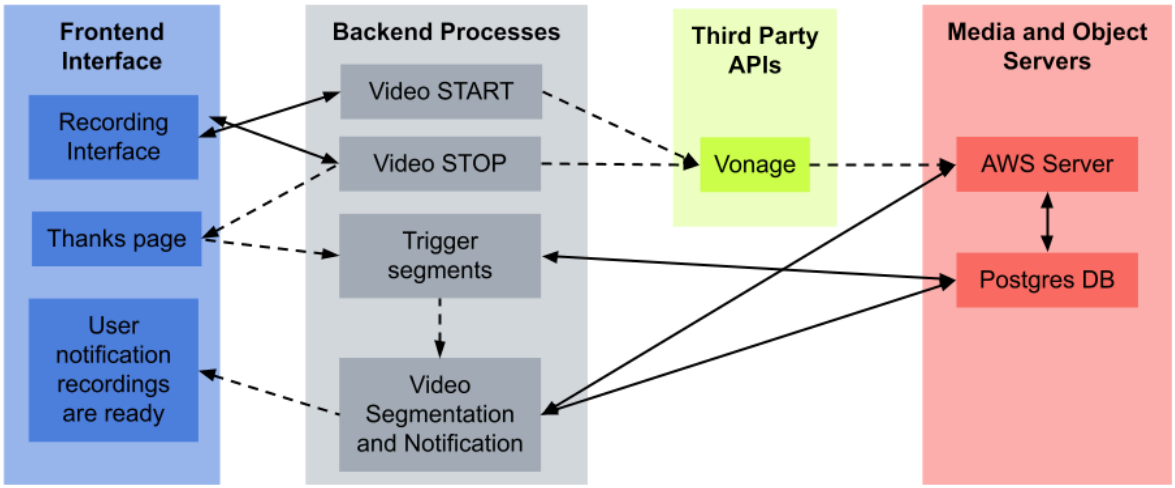


Figure: Higher level system diagram of *A Counting: Sign Language*

While there are a variety of similarities, there are some key differences, particularly in how the queuing of jobs is handled in the Signed edition and how a variety of the information is stored and transferred throughout the system.

4.3.1: Web domain and hosting

It is worth noting that the domain of sign.a-counting.us is simply a subdomain of a-counting.us, the voice edition. This is because, for the scope of this thesis and development logistics, it did not make sense for these editions to be part of the same application. Due to Heroku not allowing multiple applications to exist on the same domain [Pilot.co], sign.a-counting.us was registered and connected to the A Counting: Sign Language Heroku web application.

4.3.2: Web application architecture

Heroku hosts the Django application of A Counting: Sign Language. This is done through web containers called dynos [Heroku]. For A Counting: Sign Language there are three dynos as seen on the figure above, as well as a queue that manages tasks between these dynos and the system.

These dynos and subsequent tasks are managed by a Redis queue. Redis queues are pre-built FIFO (first in, first out) queues that are compatible with the Heroku development space. While jobs can be assigned priority levels of low, medium, or high priority and pushed to the front if needed [Redis], it is important to have the Scheduler dyno to reorder the queue as needed and to clean out failed or obsolete jobs for memory purposes.

Table: Scope of processes in different dynos

Dyno	Summary of responsibilities
Web	Handles all requests, serves all frontend web pages and endpoints needed for any processes
Worker	Handling all background tasks, including more time-intensive ones
Scheduler	The scheduler dyno runs at the beginning of each hour and is responsible for reordering and cleaning the queue, as well as initiating time-intensive video processing tasks for the video portrait.

This is particularly relevant for this project since video processes must be set at a larger timeout and memory value than default Redis queue tasks. One such example is that the portraits are all generated each hour, but in the case of a portrait that takes a bit longer, it is important to ensure that more videos are not queued up until the previous portraits are generated. The one hour scheduling and timing of the portraits themselves ensures that portraits will not be played twice, since enough portraits are generated to fill an hour of unique content by the time the portrait reloops on the web player.

Jobs within the queue are prioritized in order to ensure the video portrait generation, video segmentation for users, and credits are always updated in a way that makes the most sense for this system.

4.3.3: Database schema

There is an abstracted schema to manage the objects within the application. While it takes inspiration from the *A Counting: Spoken* edition, there are a variety of important differences between the two.

Most specifically, *A Counting: Sign Language* does not have the **City**, **Recording**, or **Merged Audio** objects. Additionally, some objects are linked in different ways. The objects and their properties are outlined below. Bolded fields represent linking attributes between objects.

Figure: Database objects and properties

Participant • hash_id • email • phone_number • created_at • updated_at • zipcode • referral	**Number** • **language** • value • original_value • invalid • url • correct_count • avg_delta • incorrect_count • **count** • thumbnail_url	**Transcription** • **participant** • numeric • written • created_at • updated_at • **number** • **count**
Video • url • archive_id • created_at • hash_id	**Comments** • **count** • **video** • transcribed • transcribed_by_signer • **transcription** • thumbnail_url	**Count** • hash_id • created_at • processed • transcribed • transcribed_by_signer • **language** • **participant** • video_1_25 • video_26_50 • video_51_75 • video_76_100 • video_comments
Language • name • hash_id • is_sign_language		

4.3.4: Administrative panel

Much of the above schema was designed with the administrative, or admin, panel in mind. The administration panel is a feature supported by DJango that provides a GUI for editing database objects, which can be registered in the panel with a variety of fields [DJango]. This panel is exceptionally important in the context of debugging and experimentation. It is also important as a piece of documentation for this code base for future developers to be able to see the objects in an easy way and create and modify test objects for future changes.

4.3.5: Information pipeline

Information travels and updates throughout our system in three ways: POST requests to different endpoints, modifications in the administrative panel, and backend modifications.

POST requests go to endpoints within the web application but also to third party APIs, such as Vonage. Endpoints of the application are utilized in both the front and backend -- with endpoints either serving information to a webpage or being utilized to trigger backend processes on the Work dyno.

Modifications in the administrative panel are only meant for testing, quality assurance, or debugging purposes. They are not planned or relied on for this system as a whole, but they are useful tools for future development of this and other works.

Since this is an object based system, most of the information that dynamically changes and subsequently travels through our system works on the object level when these objects have attributes that are modified. For example, there is never a POST request to explicitly carry the handedness of a video clip or if the video clip is invalid for the portrait based on filtering described in 4.10. This information is all handled by backend processes and by updating attributes of the video objects in the database.

4.3.6: Media Storage

The Poetic Justice Group utilizes Amazon's AWS and S3 solutions for media storage. In order to interface with the code, the S3 boto library allows our web applications to read and write to appropriate locations within our S3 storage account. Vonage also has a storage solution, however we opt for our own AWS solution.

AWS allows developers to have "buckets" (folders) connected to various applications with various policies for reading and writing of this data. These contribute to the security and organization of the media assets for the project, respectively.

AWS defines "bucket policies" in their development documents for the S3 service:

> Bucket policies provide centralized access control to buckets and objects based on a variety of conditions, including Amazon S3 operations, requesters, resources, and aspects of the request (for example, IP address). The policies are expressed in the *access policy language* and enable centralized management of permissions. The permissions attached to a bucket apply to all of the bucket's objects that are owned by the bucket owner account. Source: S3 Dev guide, Amazon

Overall, the bucket policy is a way to make sure only certain domains and actors can access information for editing and downloading of various data. This protects metadata from bad actors along with preventing bad actors from erasing or corrupting data.

Every archive video (raw stream from Vonage) is stored in a folder with the count's Hash ID. This is where the count's numbers and testimony data are stored as well. This creates an easy

system for both human and code reading. Furthermore, the credits and numbers JSONs are stored in the bucket for the project.

4.4 Security

With any system involving videos, security is a very important concern. Both video and participant information exist in the *A Counting* system but various measures are taken to ensure more security for this information.

4.4.1: Video security

While building a security system is out of the scope of this thesis, several design solutions were created in order to consider security of the video files and the participants:

1. At no point is participant information (name, zip code, contact information, etc) visible in any endpoints that are accessible to the public.
2. At no point do the video files and their metadata, such as location or device type, leave our AWS server. This is because the Vonage account is connected to our AWS server, such that the files aren't permanently stored in Vonage either.
3. A hash ID (detailed below) is used for each video and participant object, as with the voice edition, to add security in URLs.
4. By default, it is harder to download video files from our webpage. It is impossible to block downloading of content in a DOM element, as individuals can still do some scraping of the source links, but steps have been taken to make it harder, such as hiding the GUI dropdown to do so. Furthermore, our AWS bucket policy, outlined in the section above makes downloading more difficult as well.
5. But even when downloaded, the original metadata from the recording of our participant is gone, since all videos are compressed and resaved on our AWS server before being seen by participants.
6. All videos are cropped around the hand, eliminating the face and much identifying information from the frame of the final portrait.

Overall, individuals can gain access to these video files but the information a bad actor would be able to glean is only the visual of the hand performing the sign or in the case of testimonies and transcriptions, the person signing as well. However, all identifying metadata is removed and downloading the files is non negotiable, discouraging bad acting or trolling.

4.4.2: Vonage storage and security

The video stream itself travels within Vonage's network and then is written to a file saved to our S3 bucket.

Whilst this stream is being recorded, Vonage's API service provides the AES-128 encryption security to prevent bad actors from attaining these streams and their metadata [Memos, Vonage].

4.4.3: Hash usage in system information

Both editions of A Counting utilize a hash ID system for the counts and participants. These hash IDs are randomly generated by computationally salting and hashing the default ID when objects are created in DJango through the Python library hashlib [Python.org]. Thus, each video and participant has a unique 4 character identifier that is randomly generated and difficult to guess.

This makes for an easy way for users to access their information while preventing bad actors from accessing or identifying any users as long as they do not have the salt by which our hash was derived. This information is protected in the Poetic Justice Heroku organization and account.

Furthermore, only hash IDs are used for access and data transfer via endpoints -- preventing the exposure of sensitive properties of the objects.

4.5 Segmenting individual signs

As alluded to in previous sections, this project is not about the recognition or learning of various signs. However, from a count of 1-100, individual number signs must be segmented in order to be combined into a final video portrait with a different hand for every number. This mimics the audio edition, which detects silences in the audio files and slices at these silences to separate discrete numbers. Instead of slicing at silences, segmentation occurs based on the timed prompts in the recording interface where users are pausing then entering the next sign. There was experimentation that resulted in this methodology and there is still room for future improvements, outlined below.

4.5.1: Original segmentation methods

Originally, the segmentation approach was to utilize computer vision to detect an in-between sign and slice the videos appropriately at this slice. This was another reason the user testing was designed for 3 different states that were thought to be feasible for signers while also being feasible for segmentation.

Table: 3 possible in between states for segmentation

Method	Compositional influence	User experience influence	Computer vision approach
No hand between - hand leaves the box	Compositionally it would be difficult to see hands entering from different sides to make the sign	Signers felt this would be too tiring to do for all 100 signs	Very accurate since just detecting if there is and isn't a hand
Fist - making a fist at the end of each sign until the next prompt	Adds uniformity but may get cropped in or bias the composition towards closing at the end of signs	A closed fist is a sign in itself and would be very confusing within a number portrait to both sign but also if it was accidentally cropped into the portrait	Contours or mediapipe geometries
Pausing - holding the end of each sign for a few seconds	Strongest visually to hold the signs	Easiest according to user testing	Motion detection and contours with heuristics Checking for vectors moving or not moving in space

Since fist and pause were more complex to do for segmentation than the presence of hand or not hand, a fist was tested as this was a portion of the first data set that was later used in user testing. Special thanks again to Timothy for doing the first count to 100 for us with a closed fist!

OpenCV was utilized in order to perform background subtraction and find contours. Particular contour conditions (i.e. the lack of local minimums and maximums) were classified as fist or not fist for this methodology.

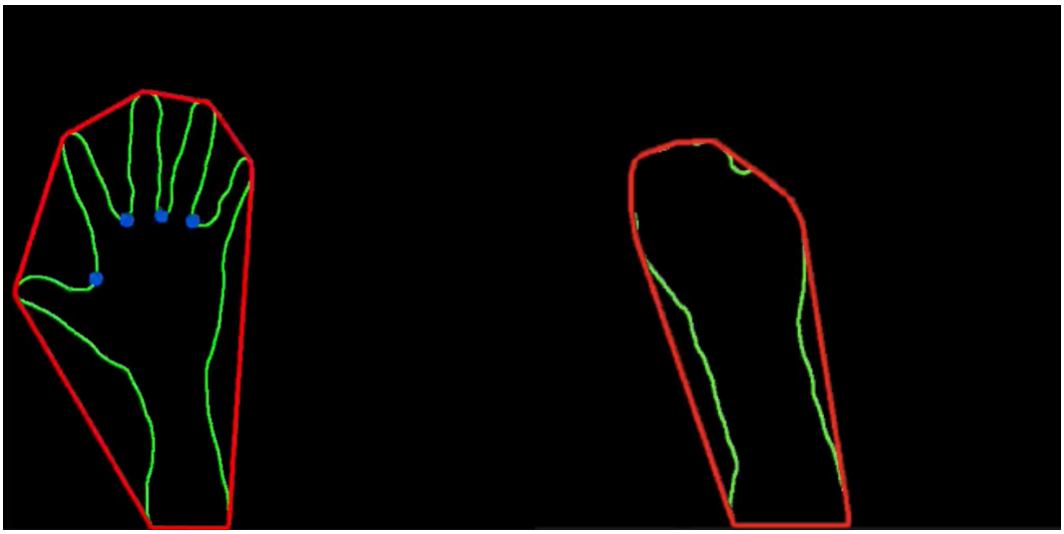


Figure: Example background subtraction and contours of a hand in multiple positions in OpenCV.

This method was fairly accurate, yielding 93 slices out of 100 numbers from our test count.

However, utilizing Mediapipe's geometries of a fist with a wrist anchor thanks to the landmarks in Mediapipe to detect a shift in the sign proved to be quite effective. This methodology was able to correctly slice all 100 numbers, helping with the fists that the contours were not able to.

At the time, this was all implemented in a Google Cloud virtual machine with mediapipe in C++. However, since then, any Mediapipe code is now being done in Python integrated into the Heroku development environment. This is further seen in section 4.6 below.

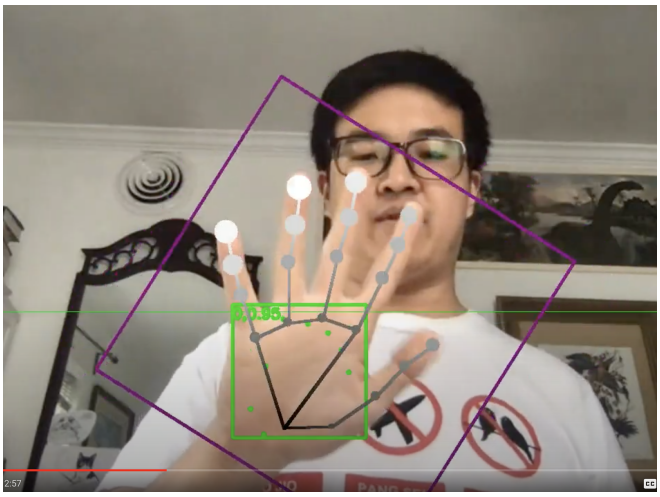


Figure: C++ Mediapipe annotations in a video sample from Tim's test count

Pausing was tested in the context of Mediapipe landmarks, which is a simple geometric interpretation for later samples. Overall it was very accurate and also resulted in all the slices. Pausing was not tested on samples that do not work with Mediapipe, however experimentation around this is highlighted in section 6.3.

4.5.2: Interface affordances and solutions

Amongst examining segmentation methods and understanding there would be inevitable errors, a simpler solution that would provide more consistency across all counts emerged thanks to affordances in the recording interface.

Since users are given timed prompts of the numbers, it was evident that a faster, simpler solution would be to slice the video clip at these transition point timings. This process is accomplished by simply slicing the video clips at these timed increments, starting at the end of the video clip to avoid any overlap at the beginning of the clip due to the Vonage API.

The above process is made possible due to the structure of the videos and interface system itself -- each count is 4 videos, and users click to start each segment. Starting from the end ensures that the slicing is not affected by any in-between time as the stream is starting and the user is entering the frame, including the countdown in the interface for the first section.

This is done with built-in functionality from the moviepy library, which takes these timestamps and slices and compresses these number clips for transcription.

4.5.3: Errors

Errors are inevitable in any process at this scale. Like the audio version, files can be corrupted, slices can go wrong, a pet could come in and distract the user, etc. No solution is perfect. Along with ownership of the language representation, this is another reason transcription exists for both editions of the work -- computers make mistakes.

With the method of slicing at the timing of the prompts, it is estimated that the system has an error rate of approximately 15% (+/- 5%) of the timing based on a randomized test upon writing (June 2021). Error rate for the purposes of this segmentation is defined as when signs are sliced in a way that includes multiple signs or cuts off the sign at an illegible point.

In general, there are 2 most common cases regarding this error rate: videos having legible overlap from the previous sign (e.g: “17”’s video having a clear 16 as well as the 17) and videos having legible overlap from the next since (e.g: “17”s video having a clear 18 as well as the 17).

The best hypothesis is that these are from participants either anticipating the next number or taking an extra moment to read then sign the next number.

Furthermore, it is worth noting that sign language states, unlike those of vocal languages, are not binary. Signers will always have to enter the next sign from the previous one, adding a transition state that just doesn't exist in audio pause, since the pause is audio in the absence of voice then voice coming back and the pause here is the hand being still then moving to the next.

This is part of the reason why there is a built-in margin within the slicing code that can be adjusted at a later point as more recordings and trends emerge.

What is also good is that we can keep changing it — overall these are cut with an increment (currently based on the prompt timer) but this can be modified in the backend or in the future based on any trends that emerge.

4.6 Centering and scaling around hands and aspect ratios

While segmenting the count into individual number videos was an important process for the final portrait, the other significant processing was to make the videos be uniformly centered around the hand(s) to result in transforms like below:

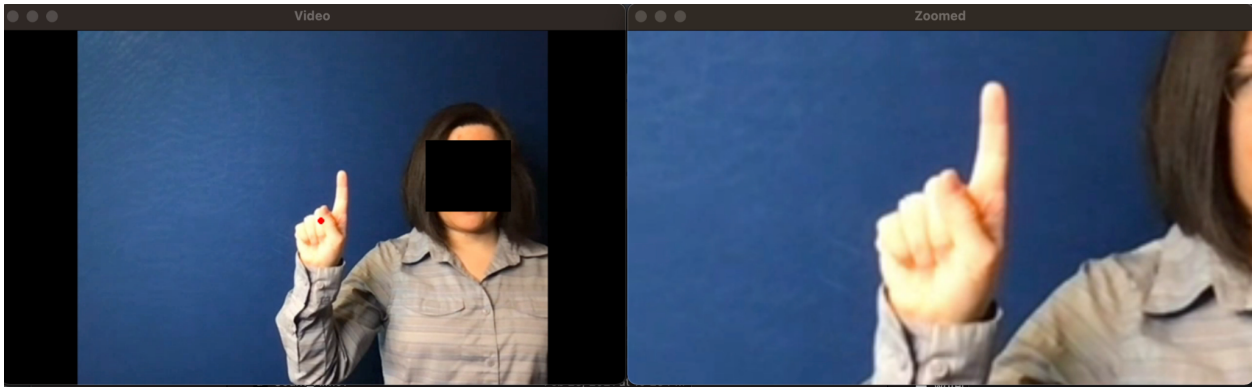


Figure: An example of the original video (left) and the zoomed and centered version (right), where the frame is centered on the found center of the hand (the red dot in the left frame). Note: face blacked out for public consumption.

4.6.1: Bounding box and center

In computer vision, there is a notion of a region of interest or bounding box whilst doing methods like object detection.

Hands are no different. Mediapipe by default identifies a bounding box around landmarks and the hands. This bounding box however is never guaranteed to be a square or the same relative size, but instead only guarantees to encompass all of the landmarks of the hand:

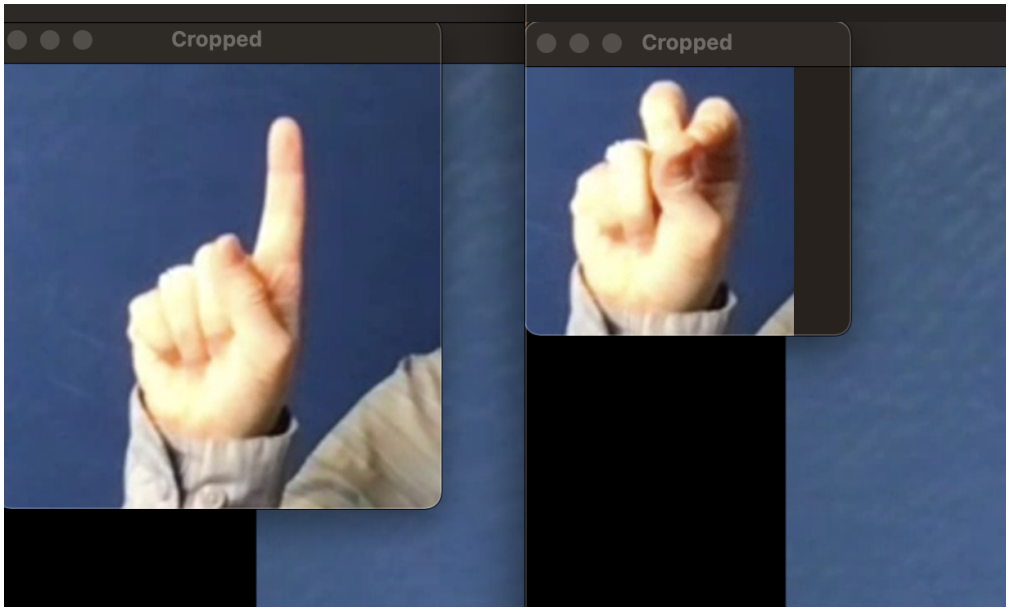


Figure: Examples of raw types of bounding boxes that Mediapipe and similar computer vision algorithms may generate, which change shape with the hand itself.

Originally, the portrait was being cropped to be a square portrait utilizing these bounding boxes, as seen below from both the raw and the transformed square bounding boxes:

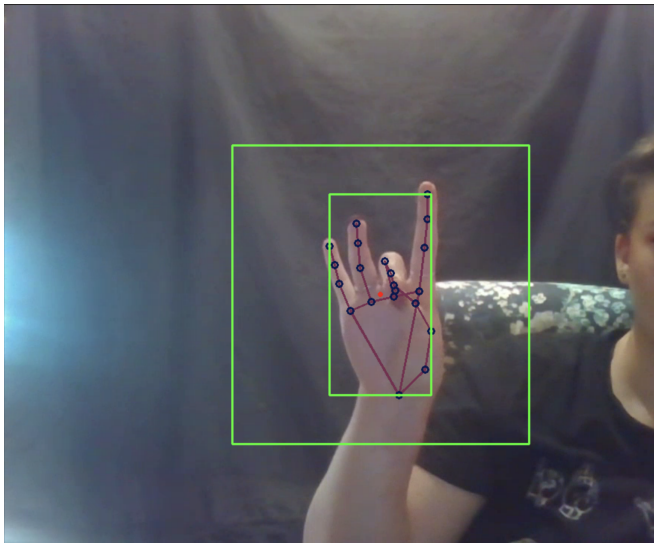


Figure: The inner bounding box of the hand from Mediapipe, with the surrounding square box with some padding for a potential square portrait.

However, the aesthetic decision was then made to make the video portrait full screen landscape aspect ratio and to allow for background and not just the hand.

But these square bounding boxes were still extremely important in order to guarantee that the center was correct and consistent. Furthermore, to ensure that even margins were guaranteed around the hand in the final crop. The center itself is simply the centroid of the bounding box, with the square bounding box an encompassing square of the automated bounding box, as seen above. It is implemented as a list of points, as the centroid moves with the hand, as seen in the figure above where the center is based off the hand.

4.6.2: Zooming at the center

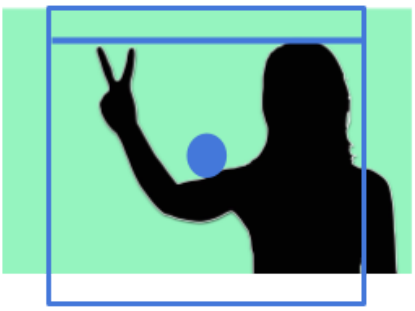
After the center of the region is identified, it is possible to transform the clip in order to zoom into a particular point. It is important to set up design requirements for such a process:

1. The bounding box's top line must fall under a top margin to ensure that there is enough compositional space between the highest tip of the hand and the top of the frame
2. The entire bounding box must fit into the frame, even at the cost of making the top margin larger
3. The center of the bounding box must be at the center of the frame, but not at the cost of the first two conditions
4. The entire 4:3 final frame must be filled, such that there are not black margins

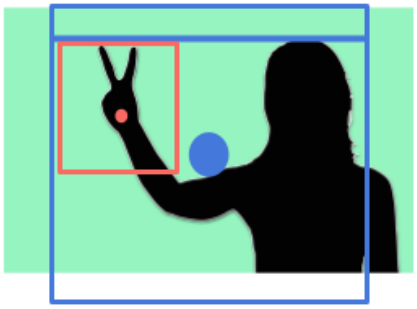
Below is the process by which raw video frames are made into zoomed in, cohesive clips of any size with respect to the user's hand. The following is done with OpenCV and Python.

1. Raw frame of user counting

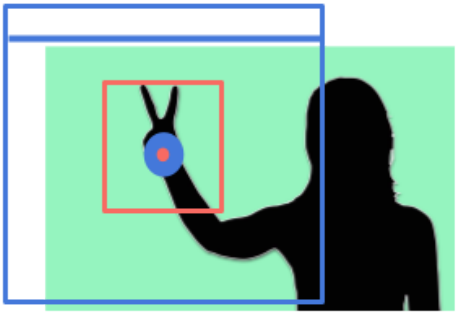
2. Identify the destination frame, its top margin, and its center and the center of the hand



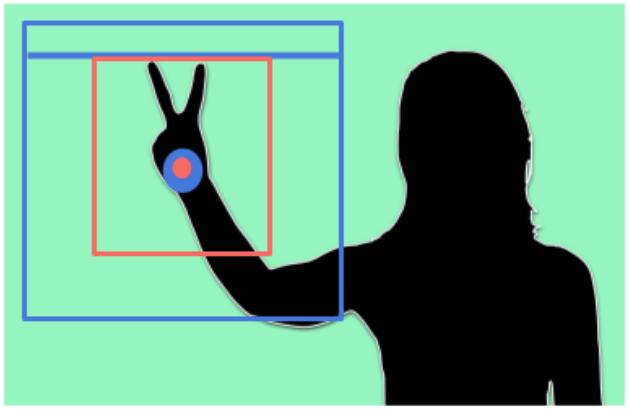
3. Identify the center of the hand and its square bounding box



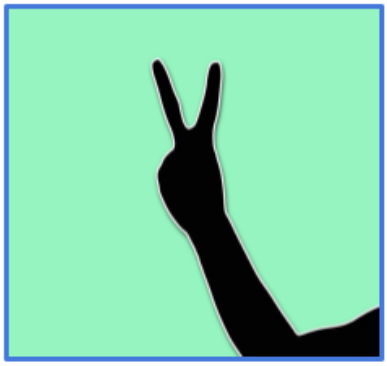
4. Move the original frame so that the hand's center and the destination center is the same



5. Increase the image size will keeping the location of the center constant until the top of the bounding box is the same as at the top margin of the frame



6. Crop to the dimensions of the destination frame



The above is accomplished by scaling the video frame based on a zoom ratio with respect to the illustrated parameters of the desired top margin, the square bounding box, and the centers. All of these are found with respect to the conditions specified above. Additionally, videos can be filtered out at this stage, as further specified in section 4.10.

Without an early exit, the clip itself must be transformed to achieve this new geometry of its zoom and XY transform. There are many methods to accomplish this, but the most efficient way is to utilize standard matrix transformations. In particular, a perspective transformation is most concise for this method. Perspective transformations are types of linear transformation matrices that take four points and map them to a new coordinate space, maintaining all of the relative points [OpenCV]. This is perfect for this case, resulting in efficient transformations once the geometric constants are found and does not require extra writing in memory or extra transformations.

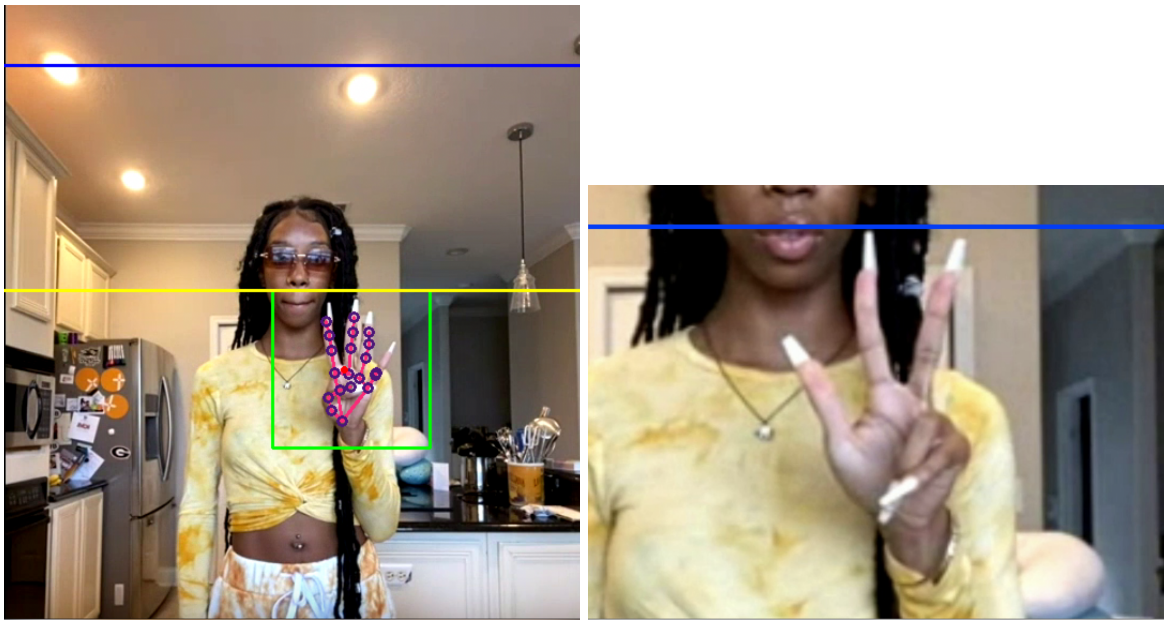


Figure: A top margin (blue) at the beginning of the original frame, along with the top of the square bounding box (yellow). After the transformation, the yellow line becomes the blue line, with appropriate zooming to maintain the entire square bounding box in the frame, centered at the hand. This is the before and after of the processes described above.

4.6.3: XY Pan

Sometimes, participants will move their hands during a sign -- either in the XY (left or right) direction or in the Z direction (moving closer or further to the camera). Additionally, some signs have horizontal motion as part of the sign, as shown below with “22”

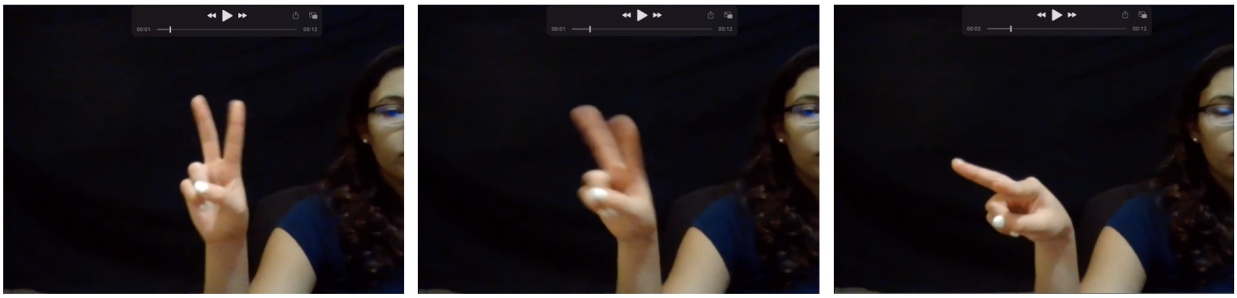


Figure: A partial frame by frame of “22”, showing left to right motion of the sign

In order to compensate for this while keeping the video portrait smooth, it is necessary to have an XYZ pan. XY and Z are handled separately. Z panning is covered as a heuristic in the previous section detailing the zoom methodology.

XY are simply changes of the centroid and can sometimes change the zoom ratio as explained above in 4.6.2. All centroid changes fall into a piecewise linear averaging equation, which is found by maintaining the coordinates in a linked list in a frame by frame manner.

This results in a methodology that any movement in the XY direction greater than a certain threshold will change the centroid sequence in a linear method. If none of this motion exists, the centroid will be a constant and not a sequence, as in section 4.6.2 showing both cases, and thus the zoom ratio handling stays consistent.

Similarly, some digits in sign language may require 2 hands. In this case, the centroid of the region of interest would be the midpoint between both hands’ centroids. The current dataset does not have such examples, so this method could not be fully tested.

4.6.4: Mediapipe error handling

There were some videos that were not able to be detected in Mediapipe. There are a variety of reasons for this, which are addressed more in depth in Section 6.

For these cases, a convex hull is utilized to detect the hand region and center upon it. This is adopted from the methodology that was originally built for segmentation. This is accomplished by thresholding the images by taking their deltas and masking them for hues of skin. After this, it is easier to identify the edges of the hands. From this, it is easier to get the centroid of the hand region, using a geometric convex hull to understand the region of interest (ROI) of the hand and gain a center of an interior square of the hand based on local minimas and an entire border of the hand.

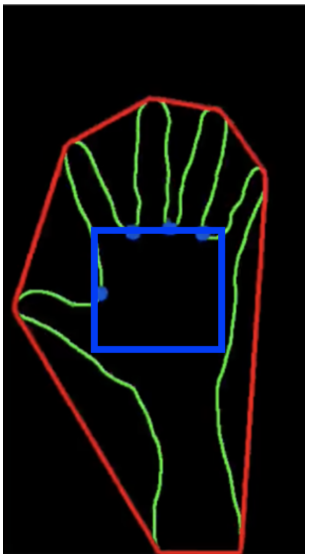


Figure: Hand detection pipeline using geometric methods, with red ROI of the detected convex hull, blue encompassing rectangle from local minimas, and yellow centroid. Styled here for display purposes. This is after the thresholding of the skin tone for the hand is done, resulting in the green outline and further detailed in later sections.

This is the centroid used for the frame centering, and it is treated the same as the found centers for zooming and panning.

4.6.5: Aspect ratios and cropping

For exhibition purposes, the final video portrait needed to be a full screen, landscape experience. Code was designed to export videos at a variety of aspect ratios by resizing the scaled and zoomed frames (as expressed above with the write-out parameters in the heuristics). A variety of ratio clips were looked at and a 4:3 aspect ratio was decided on for the final video portrait. This design process and decision is further discussed later in section 5.

Figure: An example of test frames of different aspect ratios; 4:3, 3:2, 16:10, 16:9 from left to right

By default, current camera landscapes are set to 16:9. In order to accomplish scaling and adjusting to these aspect ratios, every clip was zoomed and cropped accordingly.

4.7 Super resolution

In order to preserve the images across many different resolutions, every sample needed to be upscaled from the original HD resolution to a larger resolution. This is for gallery displays but also such that smaller resolution videos can be buffered into HD resolution portraits with larger resolution videos.

In order to do this, pixels need to be inserted at the finest detail to preserve the detail of the image. This is done by super resolution algorithms or algorithms in this family:

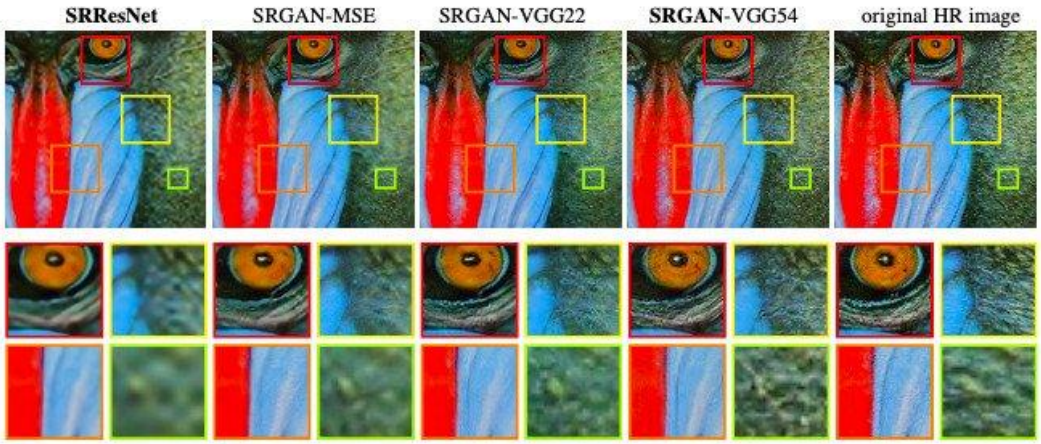


Figure 6: **SRResNet** (left: a,b), SRGAN-MSE (middle left: c,d), SRGAN-VGG2.2 (middle: e,f) and **SRGAN**-VGG54 (middle right: g,h) reconstruction results and corresponding reference HR image (right: i,j). [4× upscaling]

Figure: A figure showing the close up of super resolution on a sample. With each upscale, the photo becomes more detailed as more pixels are added with the increase in resolution. Source: [Ledig, et al, 2017].

At the highest level, these super resolution algorithms utilize multiple layers of GANs in order to create pixels that match the patterns of the surrounding area, thus replicating and adding more detail to the image to recreate it at a higher resolution [Ledig, et al]. These layers of GANs depend on pre-trained models into an infrastructure to build out these pixels with image convolution [Ledig, et al, 2017].

There are several implementations of super resolution with OpenCV and Python. These rely on premade DNN models that feed into OpenCV's implementation of a super-resolution architecture. Two models of similar accuracy according to existing literature [Ledig, et al, 2017], were tested for our use cases based off runtimes and size of the model itself:

Table: Models that were test and observations

Model	Size of model	Notes
ESPCN https://arxiv.org/pdf/1707.02921.pdf	100kB	2x, 3x, 4x available
FSRCNN https://arxiv.org/pdf/1608.00367.pdf	40kB	2x, 3x, 4x available

In this case, the FSRCNN model with a 2x resolution made the most sense for this system, mostly due to the fact that it's small file size made for faster prototyping on Heroku builds. Both models performed very well on the samples within the context of this work, especially since the samples are already HD in resolution natively, since Vonage encodes this in the stream resolution in the recording interface.

For implementation purposes, the OpenCV superresolution architecture was modified to become faster and lighter weight. This was done by making some lower level changes to memory pointers within a local version of the open-cv-contribs library, particularly with regard to how the images are stored and subsequently passed through the GAN.

4.8 Color balance

Any image and its information can be represented numerically and geometrically through matrices and vectors. Color quantization is a long studied issue in computer graphics regarding the computational representation of the colors of an image. The idea of color quantization is representing images in geometric color spaces. Color spaces are 3D geometric projects in which each color is plotted in a vector based on different space projections and color conversions.

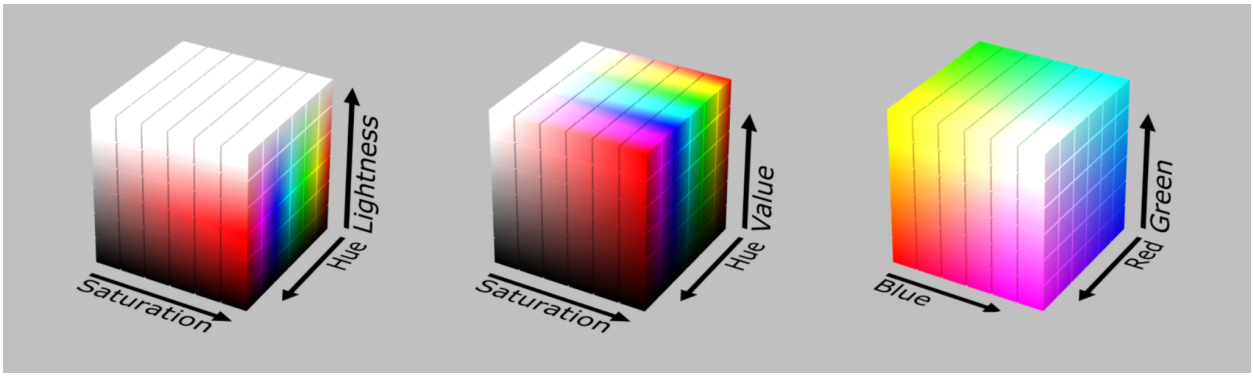


Figure: Showing different color spaces. For example, red is represented differently in RGB (red, green, blue) and HSV (hue, saturation, vibrancy) and has a different distance from the color blue. Adapted from Wikipedia.

4.8.1 Overview of color/white balancing

Color and white balancing often refer to the same process. Overall the goal of white balancing is to make neutral tones appear more neutral, often by normalizing the RGB channels to make the resulting image have less extreme temperature peaks. Within the context of photography conditions, this method tries to make an image look like it was taken under neutral lighting conditions instead of a very cool or very warm or very dim light.

White balancing is motivated by the human vision ability of color constancy based on how our eyes perceive color and then light in separate sensory organs [Foster]. Thus you are aware, under most conditions, that colors are still themselves under different lighting [Foster].

Figure: A standard example of color constancy and its applications to visual media is holding various colors and objects under different lights. While the colors and objects look different, most humans with correct color vision are able to understand that they are the same objects and colors, just under different illuminations. This constant of these objects is where color constancy gets its name. Source: Lotte Krull via DTU.

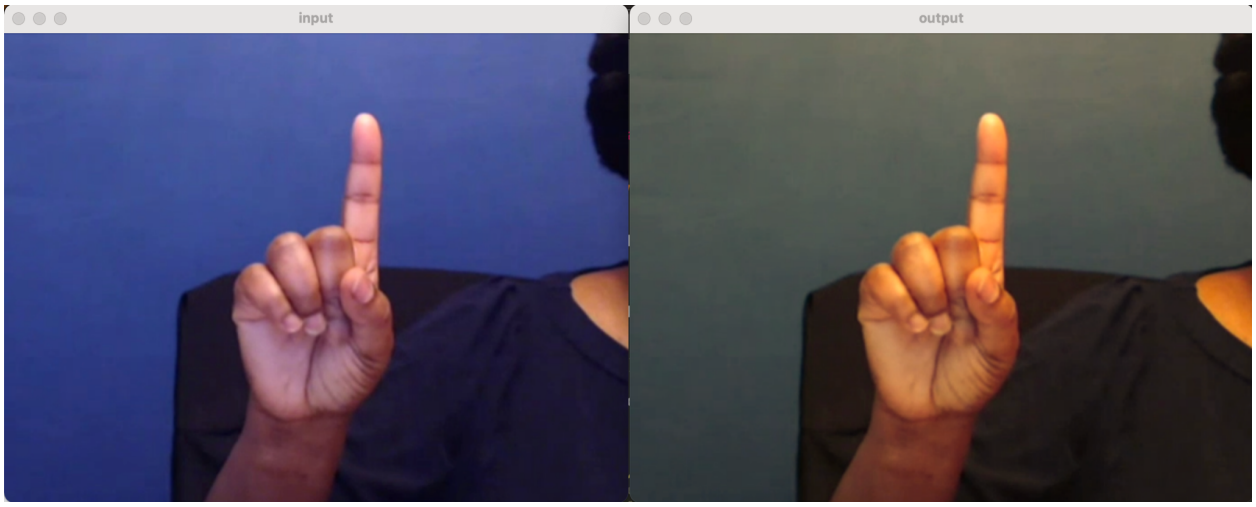
Cameras try to copy this by exaggerating exposure and color temperature, often causing noise in the RGB spectra which can result in colors not looking as they do in real life. This goes without saying the biases under which these cameras were developed and the many efforts going on to make cameras that better capture an accurate image of more subjects [Lewis, 2019], which is discussed in different aspects of this thesis.

White balancing seeks to undo these extremes, through several different methods -- most commonly and efficiently through a histogram equalization, which seeks to flatten the color curve and to reduce spectra spacing between global and local maximas. Other methods include color normalization, contrast stretching, and more.

Overall, digital image processing -- from camera development to editing -- has extraordinarily racist and colorist histories. This history lives on with respect to a variety of color balancing and other photo editing algorithms, which often cast blue and green tones across whole images in response to noise in backgrounds and to balance out red values, as seen below.

Figure: Example of a simple white balancing algorithm creating extreme blue casts across multiple skin tones. This implementation is covered more in depth below.

This is evident in darker skin tones even in the most recent Adobe Research paper [Adobe, 2019], which resulted in very bad results even when normalizing for the noisy background:



And in the most recent Samsung Research color balancing paper [Samsung, 2020], which uses a deep learning set to do automatic white balancing with respect to various lighting conditions:

Overall, it is obvious that a diversity of skin tones was not the concern in these studies, especially as a naive white balance histogram algorithm performed better subjectively and based off normalization of the resulting RGB spectra produced:

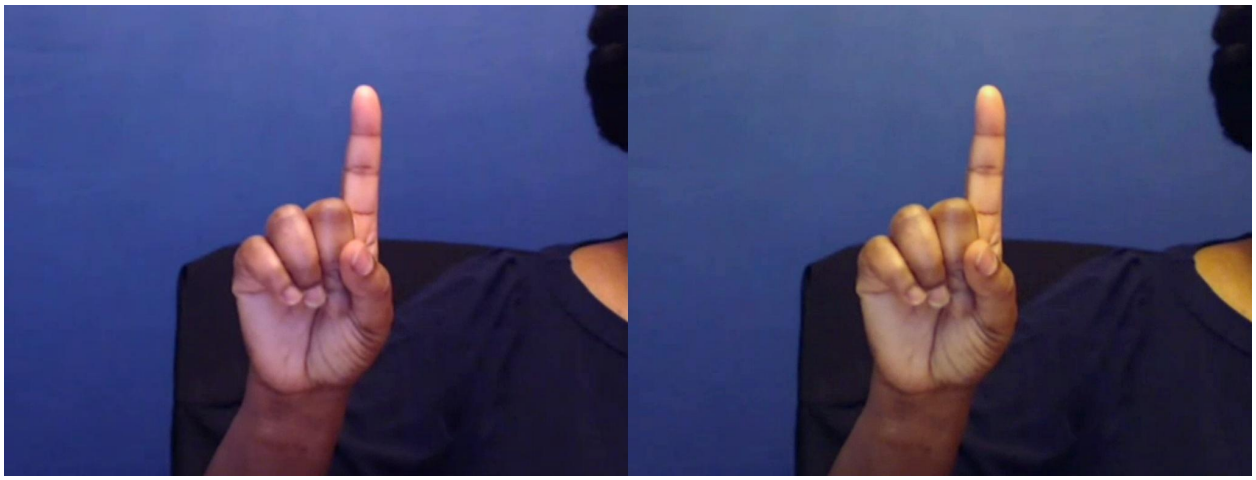


Figure: Naive white balance algorithm results.

The algorithm above utilized a histogram normalization method that is well documented in a variety of literature, such as the above studies from Adobe and Samsung. This is an extremely common approach to color balancing as it runs fairly efficiently and is straightforward to implement. It simply uses image convolution to expand the color channels of RGB into the full range of 0 to 255, thus scaling the image to encompass the entire spectrum based off a defined white point.

4.8.2: Improved thesis approach

The approach in this thesis relies on a twofold consideration that much white balancing work does not touch upon:

1. Considering how skintones are better handled in different color spaces and
2. How a histogram methodology may recreate and even stress exposure issues just in a neutralized color space.

The first consideration -- skin tones -- looks towards how the hues of skin are better quantized in HSV instead of RGB spaces. If one considers a range of human skin tones, one may consider something like this:

Figure: Rihanna's Fenty shade range, showing 50 different foundation colors on 4 models. Source: Sephora.

However, in RGB space, the colors above have this tendency of distribution, pushing darker skin tones very close to the red access and thus creating too much of a blue shift in an attempt to neutralize in RGB space:

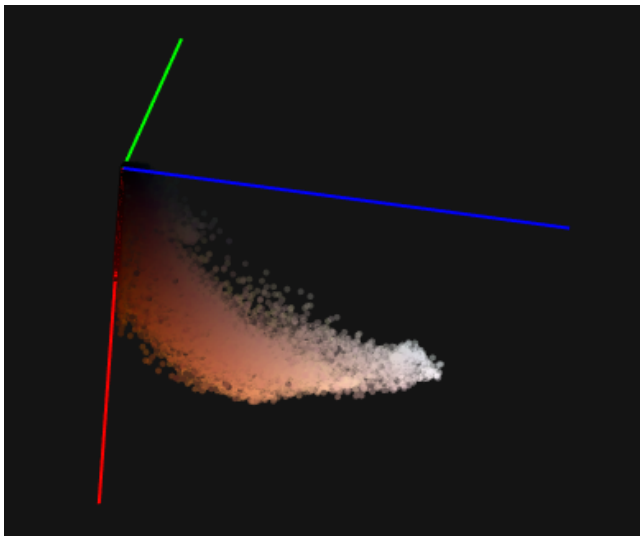


Figure: Skintones from the Fenty image projected in RGB space, showing larger clusters of darker skin tones very close together and to the red axis, while lighter and more neutral skin tones have the benefit of being more in the middle of the axes of colors, thus removing the blue tone issues shown in RGB white balancing processes.

Logically one would want a colorspace that clusters skin colors together, distributed in the same place, instead of this diagonal stratification that leads to over-shifting in editing of dark skin tones. This remains true even with backgrounds, as seen in other samples, thus much mistreatment of skin tones in color balancing can be easily avoided by leveraging an HSV or HSL color space to achieve more intuitive geometry [Lutz, 2019].

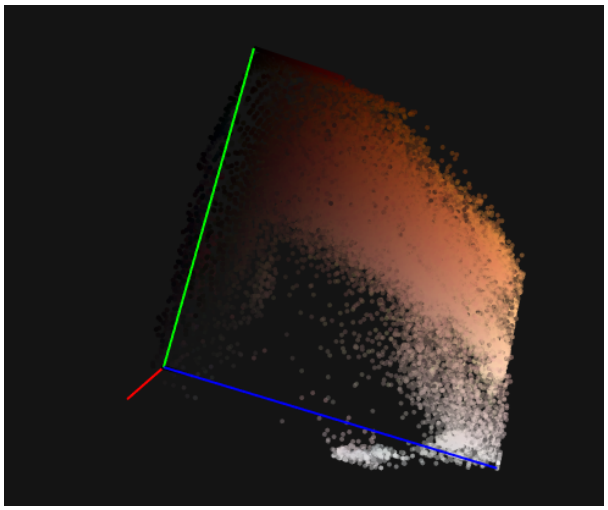


Figure: Skintones from the Fenty image projected in HSV space, showing all skin tones on the same plane, which would provide for proportionate shifting of colors in white balancing.

This type of approach results in better automatic white balancing without a blue or green cast, since skin hues are divorced from those in different geometric planes:

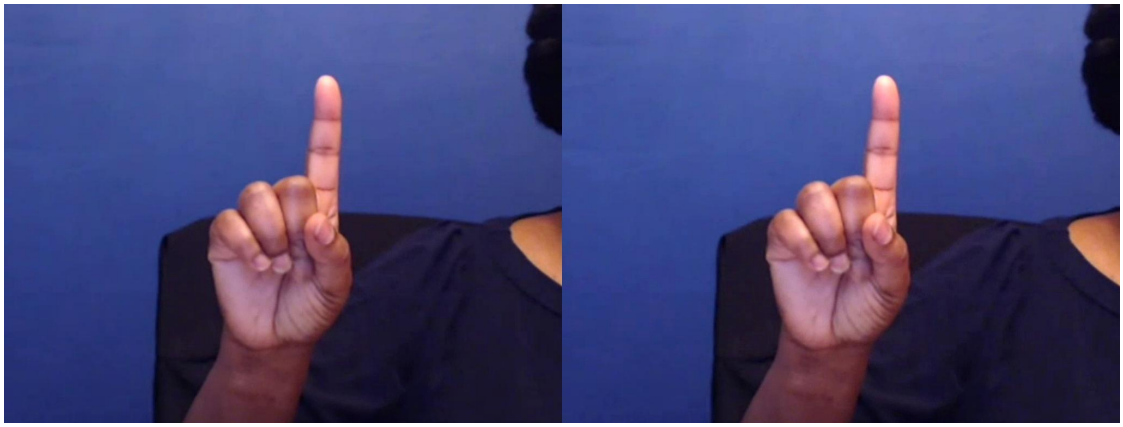


Figure: Difference of sample with white balancing in HSV, resulting in much more reasonable results.

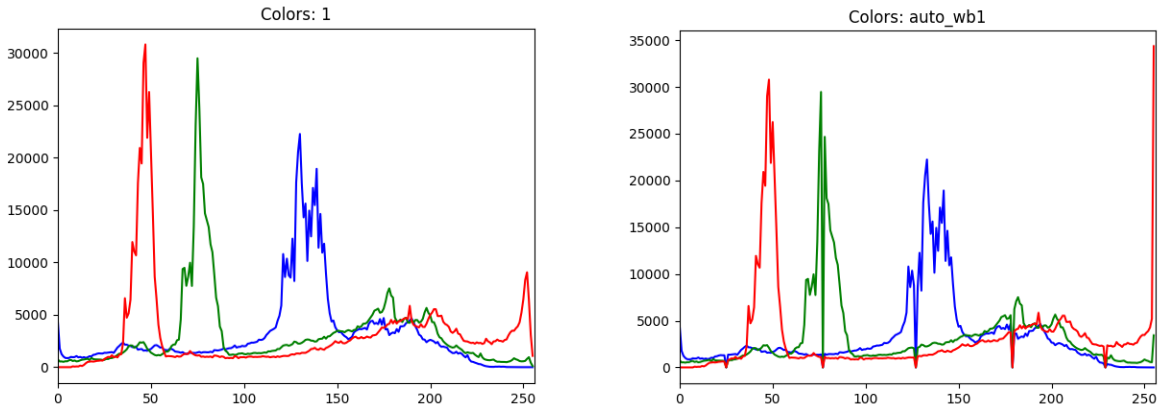


Figure: Difference of sample spectrum before and after auto white balancing with this method, showing large spikes of red in the white balanced spectra (right), but keeping more uniformity with the original color spectra (left).

Plots above show that the spectrum still stays relatively the same but becomes more concentrated in a deliberate way, reducing noise and providing more uniformity across the RGB channels in the image.

However, the second consideration remains -- this histogram style tends toward adding an extremeness of exposures, as seen in the spikes of all 3 channel spectra at the end of the auto white balance graph. This can be better seen in the sample below, which had a very extreme delta of its exposure along with color correction.

Figure: Showing exposure and color correction with thesis method

While this is desired to an extent, as the goal of white balance is to bring out and normalize amongst the neutral white colors of the histogram, this thesis desires to have a lighter editing touch on the video clips. Thus, a beta normalization version of this was developed in order to counteract the extremes of lighting and darkening of the color corrected frames, seen below in the visual samples.

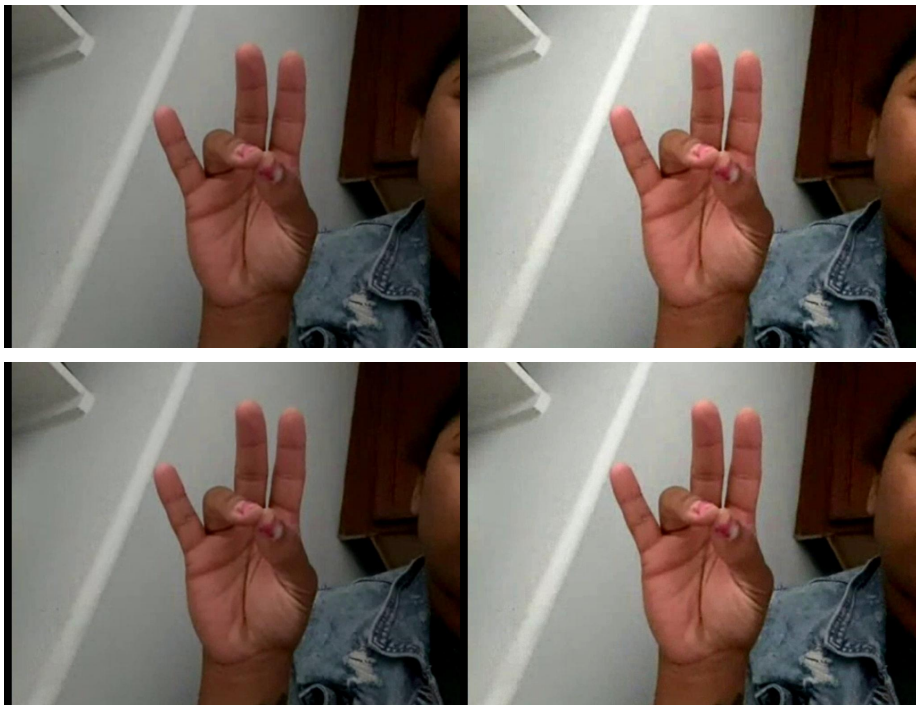


Figure: Top: AWB, Bottom: Beta normalization

While subtle visually, in the spectra plots the results of this normalization are more apparent in the spectra outputs.

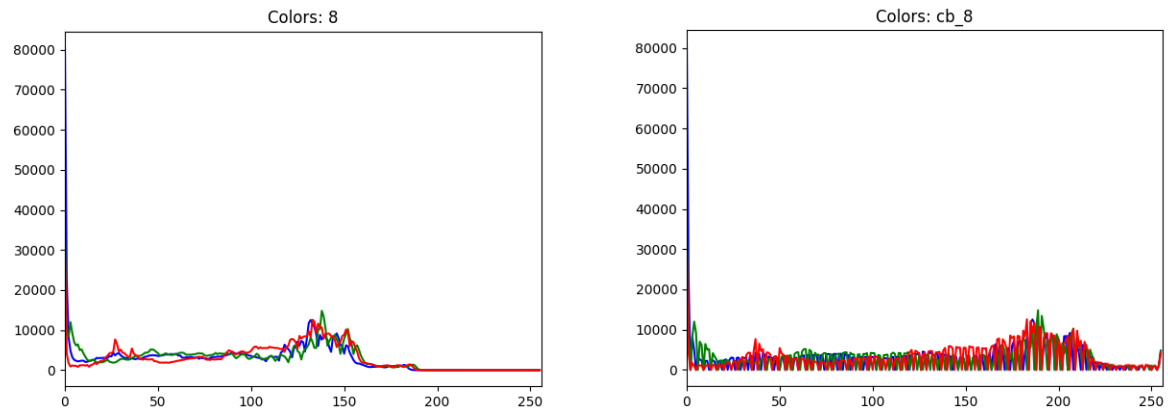


Figure: The spectra of the original, underexposed frame (left), and of the beta normalized frame (right), which added more detail and peaks in each color channel, thus adding a more neutral white to the entire image. Additionally, beta normalization still kept the global maximum for the frame at the same level, thus the same shape of the overall graphs, preventing the edited image from appearing over-exposed.

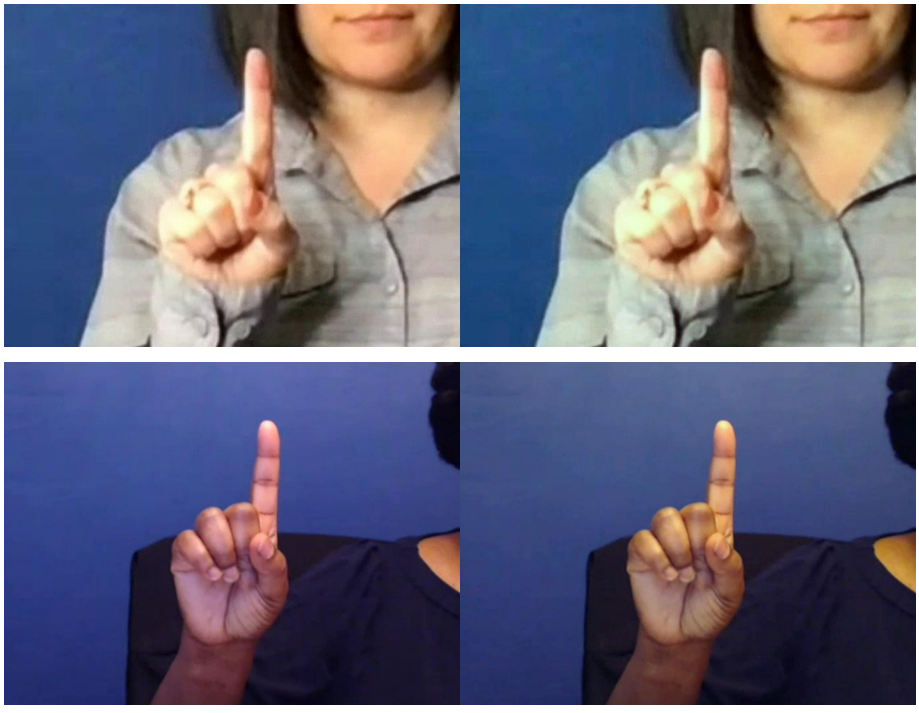
For the purposes of this document, this methodology is being called **beta normalization**, because it utilizes an upper and lower bound to normalize the histogram and prevent over and under exposure peaks. This is represented in the equation below, showing its convolution with the histogram method from above.

4.8.3: Experimentation of other methods

While developing and supporting this hypothesis, a variety of white balancing methods from current literature and softwares were tried. These results are summarized below.

Simple Color Balance: Uses histogram in RGB + grey world assumption

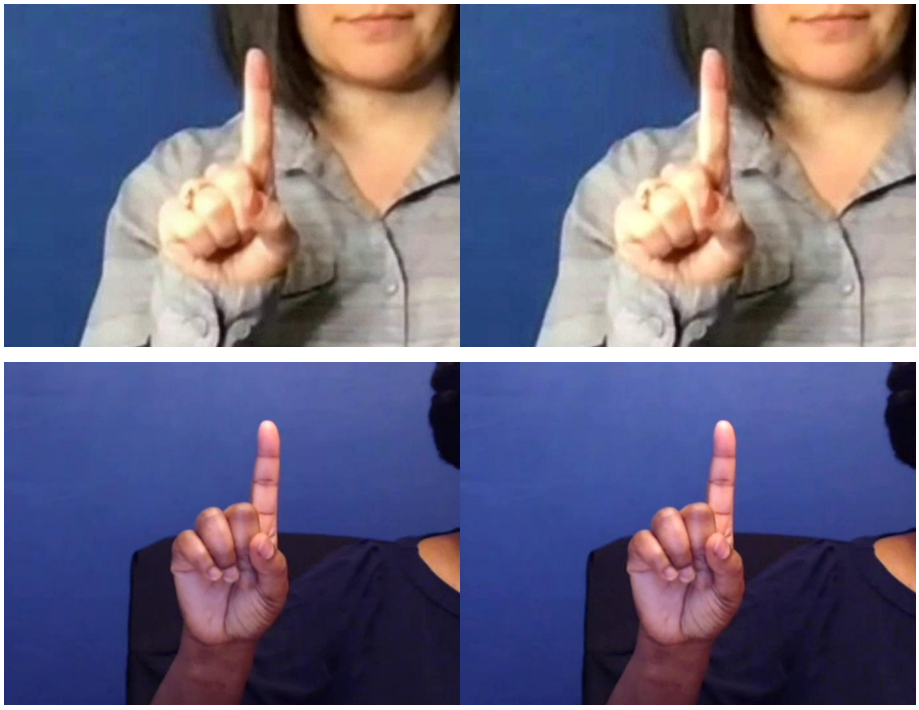
Simple White Balance: Uses histogram in RGB



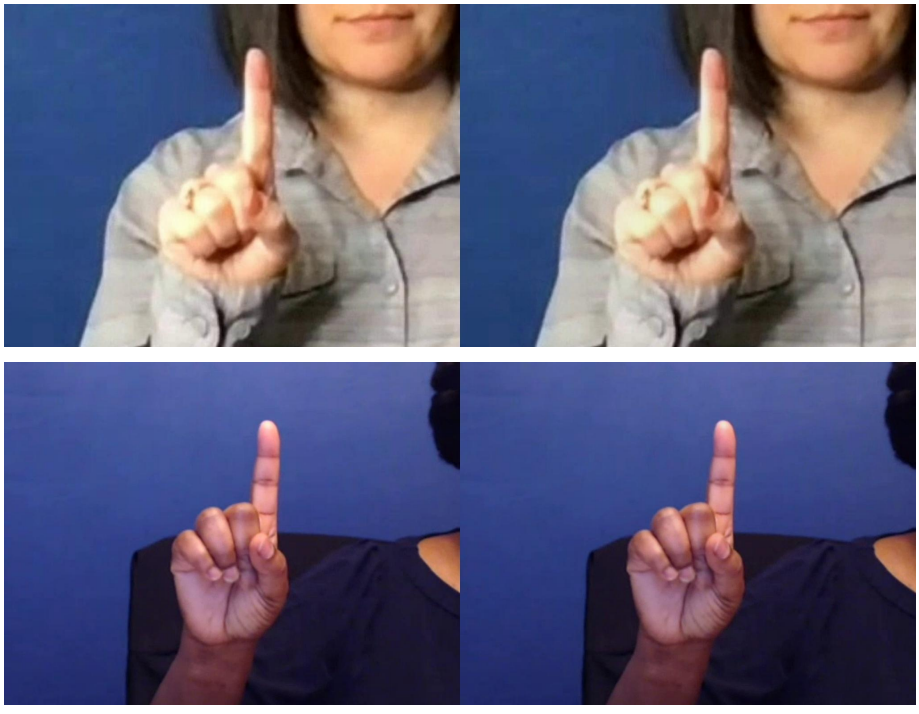
GIMP White Balance: An implementation of the GIMP editing sofware's auto white balance algorithm, based on a RGB project, that discards pixels at the end of the color channel histogram [GIMP, 2019]

Contrast Stretch: Normalizes across the darkest and lightest of pixels to expand the current pixels to the entire range in the image

Automatic White Balance: Uses histogram in HSV and normalization



Beta Normalization: Uses histogram stretching while normalizing for shadows and lights



The conclusion of these experiments is to implement a beta normalization methodology to hold a light touch of color and light correction on a variety of images and skin tones.

Overall, this line of inquiry seeks to understand and display that we cannot undo the quantification of the human form that has occurred throughout media arts history, but there are more responsible ways to establish how we enact diverse visual representation, and image processing and color quantification are just one of them.

4.9 Blur, Sharpen

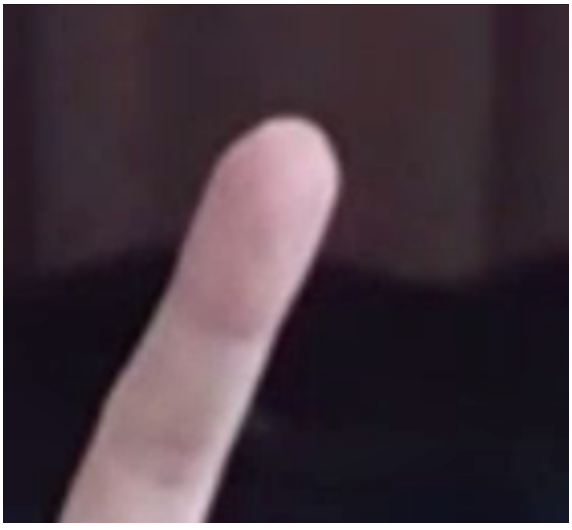
While not utilized in the final portrait, several experiments were done regarding blurring and sharpening to try to resolve the differences between quality of films.

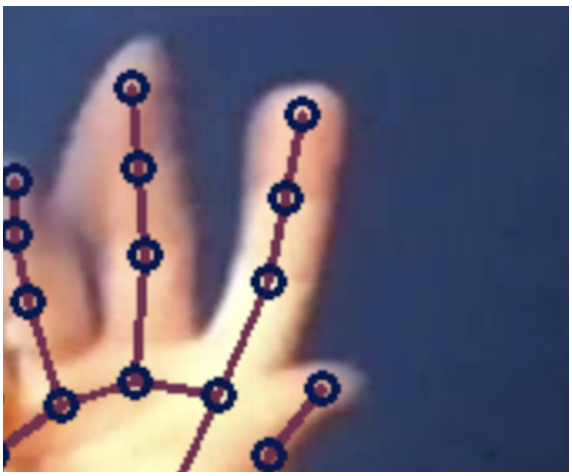
4.9.1: Motivation

Upscaling, as detailed in section 4.8, only replicates an image to a higher resolution so that an image can be viewed more clearly at these resolutions. Therefore, any visual noise from the original image will be replicated -- the super resolution will not reduce or smooth any noise or blurriness.

Admittedly, visual errors are completely subjective -- noise and lower resolution can be artistic choices and give a home video quality, which is preferred by many. However, for the purposes of this thesis, tests to reduce the most common visual inconsistencies in some of the samples were conducted. Visual inconsistencies are a problem because they appear in some samples but not others. These inconsistencies are summarized in the table below:

Table: Visual inconsistencies

Visual Inconsistency	**Cause**	**Example**
Visual noise	Different recording devices Lighting conditions	

Pixelation	Different recording devices Subjects standing further from the camera, and then scaled and zoomed in	
Blurriness and glare	Under or over exposure Different recording devices	

In order to change these inconsistencies, automated filtering would be needed. In particular, smoothing of noise (aka blurring) along with sharpening to decrease the appearance of glare and pixelation.

4.9.2: Blurring experimentation

Overall four blurring techniques were used to blur the video frames. These methods were chosen due to being widely used and documented.

Table: Blurring algorithms tested

Method	Description
Averaging	Overall normalization
Gaussian blurring	Blurring based off noise removal
Bilateral filtering	Noise removal while prioritizing edges
Median blurring	Uses the center value in each area, highly effective for salt and pepper noise

It is important to note that there is no one algorithmic process for these blurring methods -- they can be implemented in a variety of ways and new implementations are constantly occurring and have different efficiencies in different hardware and software ecosystems.

Each of these algorithms is implemented as image convolutions with different kernels. This simply takes a binary logic matrix (kernel) and goes pixel by pixel, taking the average around the kernel area, and replacing the central element (current pixel) with the result [OpenCV Docs]. This adds whatever chosen effect (defined by the kernel) across the entire image.

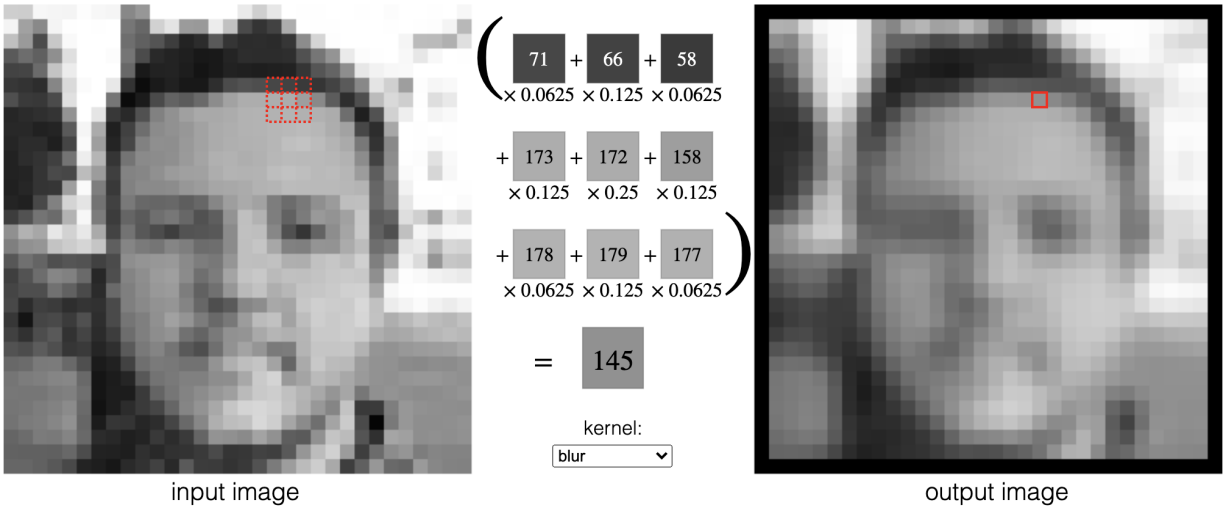


Figure: Example of an image kernel matrix being applied at the pixel level. Source: Powell, 2015

All kernel are simply defined as a transform matrix with a particular constant, which can vary or be the same constant for the entire kernel.

$$K = \frac{1}{25}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Example kernel equation from OpenCV docs showing a kernel with a k = 25 applied to a uniform matrix for transformation. This differs largely from the above kernel which varied based off an equation of contrast.

Averaging is perhaps the most basic of image blurring methods. Averaging is accomplished by convolving an image with a normalized box filter as a kernel matrix. As such, a simple use of openCV's built-in `cv.blur()` function was used to try this method for this experiment. Mathematically, this is equivalent to the above matrix, which any positive integer kernel size [OpenCV Docs]

An example result of samples with average blurring is below:

Figure: Average blurring on different skintones compared to original (left)

Gaussian blur is an extraordinarily common image blurring and averaging technique, used in many computer graphics programs. Instead of a box filter, the image is convolved through a Gaussian function as a linear transform [Perlin, 2001]. Thus, Gaussian blurring is highly effective in removing Gaussian noise from an image, which is a good approximation for a variety of noise. This is mathematically represented below:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Figure: Standard Gaussian equation. Source: Perlin, 2001

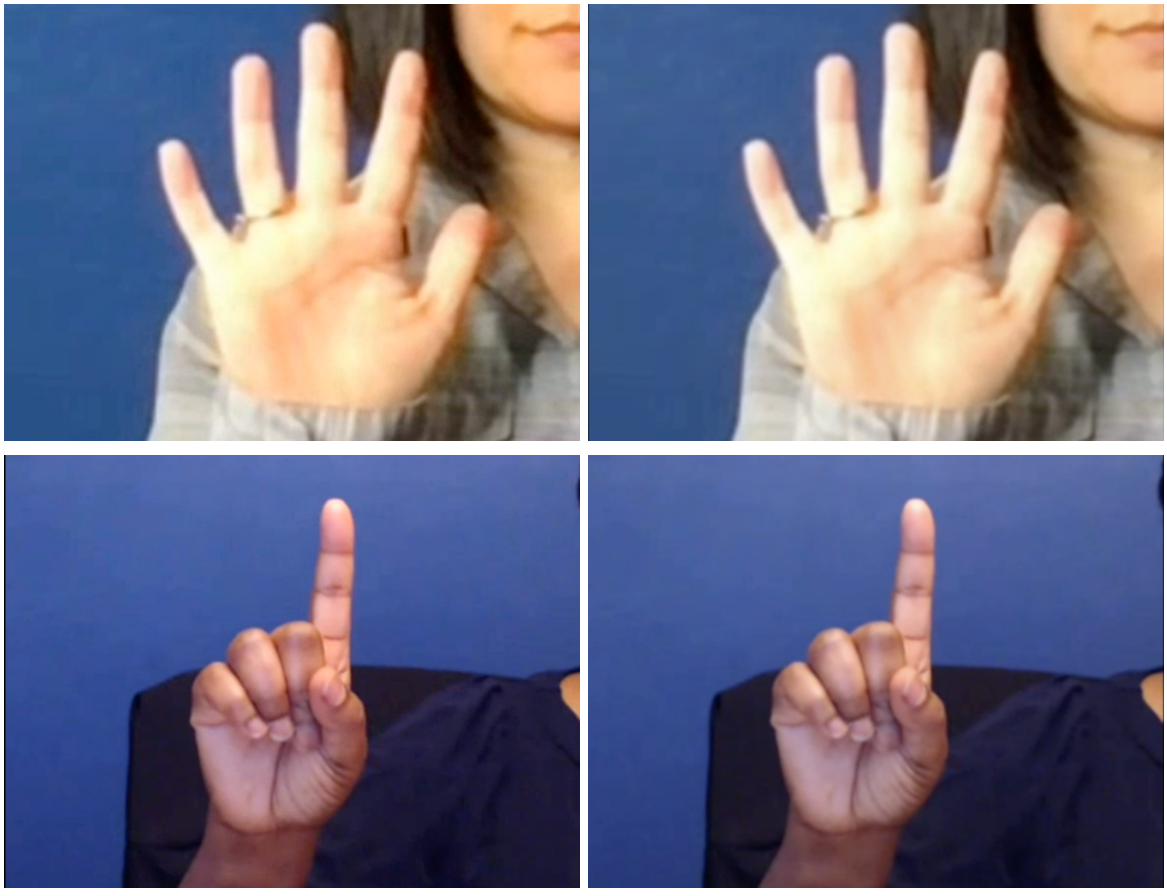


Figure: Gaussian blurring on different skintones compared to original (left)

Bilateral filtering removes noise while prioritizing maintaining sharp edges. The operation is slower because it utilizes two Guassian filters -- the first to consider nearby pixels for blurring of noise and the second to consider the intensity difference between pixels, ensuring only pixels with similar intensities to the central pixel are considered for blurring -- this preserves sharp edges, as edges are present with pixels with greater intensity variations.

$$I^{\text{filtered}}(x) = \frac{1}{W_p} \sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|),$$

and normalization term, W_p, is defined as

$$W_p = \sum_{x_i \in \Omega} f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|)$$

where

I^{filtered} is the filtered image;
I is the original input image to be filtered;
x are the coordinates of the current pixel to be filtered;
Ω is the window centered in x, so $x_i \in \Omega$ is another pixel;
f_r is the range kernel for smoothing differences in intensities (this function can be a Gaussian function).
g_s is the spatial (or domain) kernel for smoothing differences in coordinates (this function can be a Gaussian function).

Figure: The equations for a standard bilateral filtering implementation. Source: Tomasi, 1998

Figure: Bilateral blurring on different skintones compared to original (left)

Median blurring utilizes the center value in each area in an averaging procedure. It is slower than averaging, due to this need of carrying these averages throughout this linear process, but highly effective for structural noise since it centralizes all of the lightness values that already exist -- ensuring a smoother blend.

Figure: Median blurring on different skintones compared to original (left)

4.9.3: Sharpening technique tried

For experimentation purposes, a simple sharpening implementation was done with an image convolution with a sharpening kernel as seen in a variety of literature [Scarff, 1981]. This kernel prioritizes differences in the adjacent pixels, creating darker and sharper edges through contrast emphasis, as seen below:

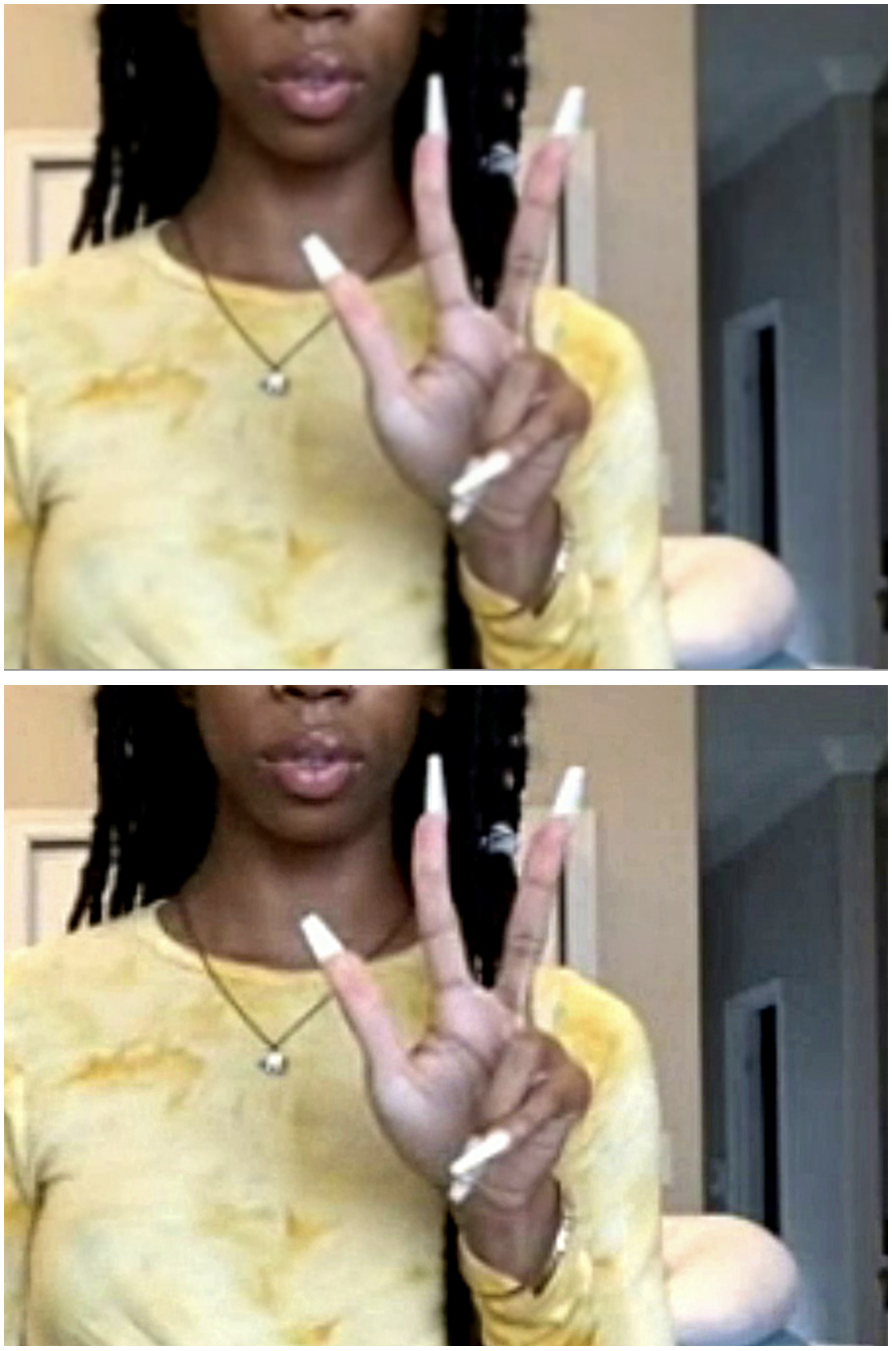


Figure: Image and image with only sharpening (top to bottom)

It was decided that sharpening created too much of an unnatural effect on many of the samples and took away from the home video quality of the original footage.

4.9.4: Results and design decisions

Overall, these differences were not consistent enough across frames and with our limited dataset to make a concrete suggestion. Furthermore, it was decided to maintain the authenticity of the video portrait that blurring and sharpening should not occur.

4.10 Final video portrait

4.10.1: Assembling the video portrait

At the time of writing, numbers from 6 counts have been transcribed by participants within the artwork.

Figure: Zoomed out frames from 6 current portrait iterations. Unfortunately, due to relatively few numbers being transcribed at this time, there is less diversity in these frames.

The video portrait assembly, at the highest level, is described in the diagram below:

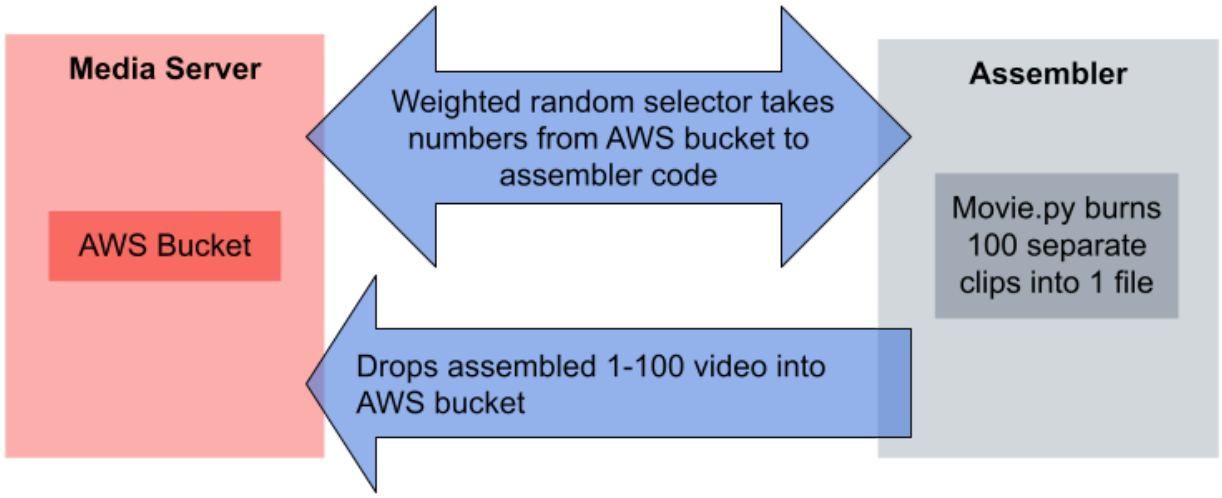


Figure: The portrait assembly process, which repeats for both the portrait and landscape files.

The first step is generating random orders of numbers that have been transcribed by participants. This is done by looking up, for each number 1-100, number objects with a correct count in the database, and ensuring the correct count is greater than the incorrect count, if applicable, and that they are not marked invalid. In the case of multiple valid number objects for a particular number, these objects are chosen with the weighted random choice operator in Python's random library based on the weight of the language of that object and the amount of correct numbers in that object's count.

Like the vocal edition, the code inversely weighs the probability based on language (in the case of multiple signed languages), and if counts have been partially transcribed. This means if there are 50 numbers of German Sign Language and 200 numbers of American Sign Language, German Sign Language is 4 times more likely to appear in each number than American Sign Language. This is to maintain variety in the portrait and to hold space for rarer languages.

Furthermore, since the signed edition will be dominated by ASL, if there are counts that are partially transcribed or have fewer correct numbers, this is also taken into account. For example, if 2 counts have 1-100 but the third count only has numbers: 4, 28, 58, 61 -- those numbers will be prioritized in the portrait over the numbers from fully correct counts. This is to add variety to the visual portrait itself.

After numbers are selected, they go into the pipeline of being scaled and cropped to the appropriate aspect ratios for landscape and portrait orientation display. The details of these methods are outlined in section 4.7.

After numbers are scaled and cropped, both the landscape and portrait orientation clips go through a filter layer to prevent numbers that have too much face after they are scaled and cropped. This checks for faces because we do not want to show faces in the final portrait to anonymize participant identities as per our terms of use and to keep the focus on a portrait of hands.

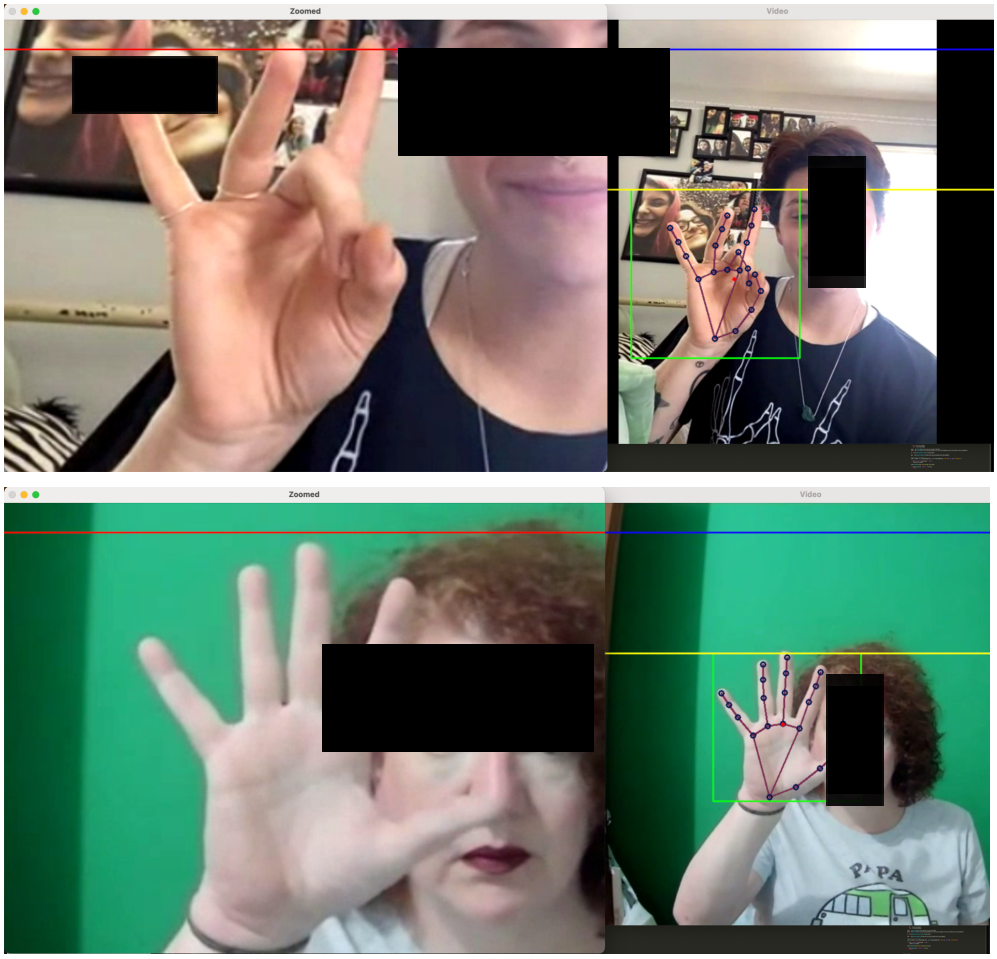


Figure: Showing the scaling/cropping of two samples where after zooming and cropping from raw footage to the 4:3 zoom for the final portrait, there is still too much face present. Lines for internal processing purposes and debugging. Note: black boxes added to cover identifying facial features for public consumption

Face filtering is done by seeing if more than half of a Mediapipe face mesh is detected.

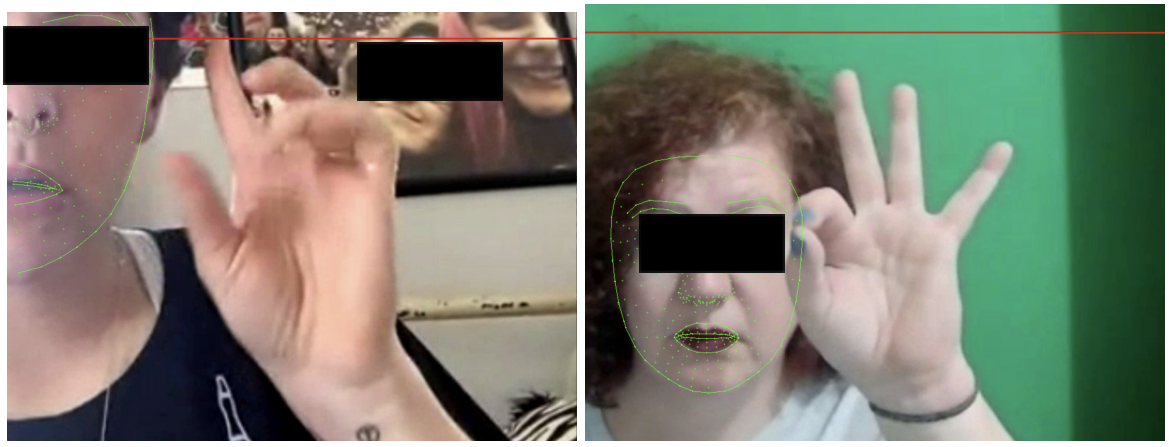


Figure: showing half and full face meshes, both of which would be cut from the final portrait. Lines for internal processing purposes and debugging.

Furthermore, after seeing cases like above, a maximum number of faces is set to be 5 in order to protect the privacy of anyone in the background, even if enough of the face of the subject is cropped out:

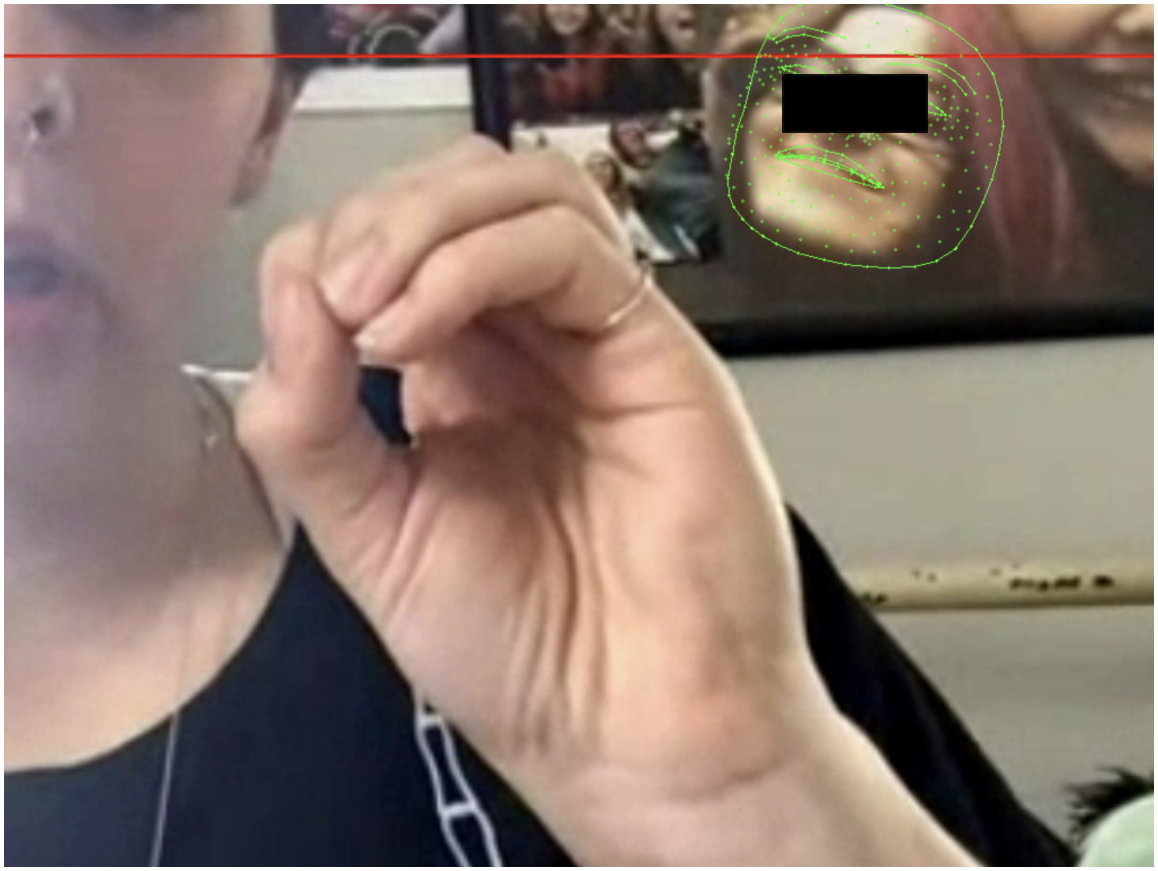


Figure showing a background face mesh. Lines for internal processing purposes and debugging.

Face detection itself has many issues as it can fail to detect faces, particularly for darker complexions, in over or underexposed frames, or for individuals wearing glasses and other face wear. This has been shown in a variety of case studies, perhaps most relevantly in Joy Buolamwini's masters thesis showing that her face was not detected by a standard face mesh algorithm, which lead to the formation of her efforts in the Algorithmic Justice League seeking more equitable uses and development of such systems.

And while Mediapipe is fairly accurate and has not had issue with our current dataset, including dark complexions, we have added another layer of convex hull code and skin masking to detect any half ellipse convex hull regions that match the skin color of the hand, and discount these as faces as well. This is done with the methods described in section 4.7.

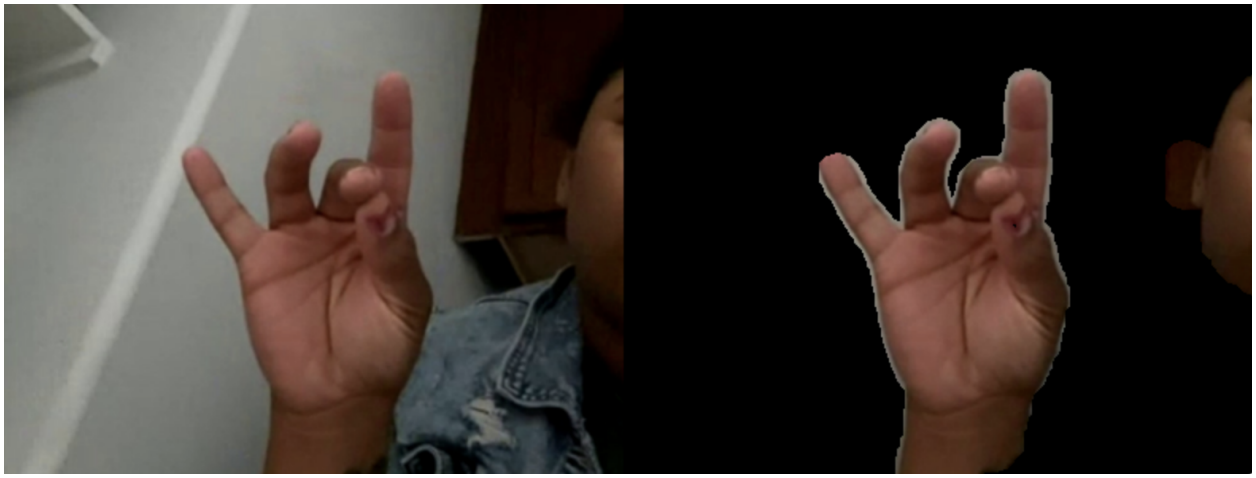


Figure: Showing approximate skin masking on a darker complex recorded in underexposed conditions

Figure: Showing masking output, hull lines, and the enclosing ellipse, which the hull lines do not fit into half of (denoted by the red hull and pixels fitting within half the generated enclosing ellipse. Resized for illustrative purposes.

If any numbers are filtered out due to face detection, that number is marked invalid in our database, which means it will not be picked for future portraits. A valid replacement is then chosen from an endpoint. These are marked for both orientations.

For any numbers not filtered out, they are marked as valid for future portrait use and so that computation will not need to be used, scaling and cropping them in the future.

These numbers are then buffered together into one video from 1-100. There is one for landscape and one for portrait orientation of 4:3 and 9:16 aspect ratios, respectively. This is done with moviepy on a scheduler dyno and saved to our S3 buckets.

This all occurs in the Heroku environment on the scheduler dyno and Redis queue. The current system is designed to generate 15 video portraits each hour to cover a unique hour of footage, since each portrait is approximately 4 minutes long total.

4.10.2: Streaming the video portrait

Just as recording a stream of video across the web is non-negligible, it is also a non-negligible task to stream a large video file(s) -- as seen in section 4.2, regarding difficulties having video files within the transcription interface.

In order to make sure that the video portrait would load in a timely manner, it is important to remember that approximately 1.25 MB/second could be loaded in a straightforward way on slower connections. Since each portrait is approximately 100 MB and 4 minutes long. Without any streaming code, that means on many networks a single video portrait could take almost two minutes -- far too long for a load time for a webpage.

Furthermore, it was important to not load a large amount of videos, such as loading 100 individual number videos or even all 15 videos to cover an hour’s worth of portrait. This is to prevent memory leak issues that slow down or time out the page.

Instead, two video objects are used and Javascript code changes their source attributes. Two video objects are held at a time in the case of network errors so that the second video has time to load while the first is playing. These source attributes come from a JSON that contains the video link for the landscape and portrait orientation views along with the information for the credits. These source attributes are added to the video elements which load quickly due to being streamed by a publisher method from the Vonage Video API that is utilized in the recording interface. This essentially allows video files to be loaded much faster from a dedicated server instead of directly from a request to the page from AWS.

Since galleries may have these slower connections, it was exceptionally important to have a fast load time. The goal load time is 5 seconds on most networks and for this video portrait to have a similar load time to the audio edition. This was accomplished, as seen by the information below from measurements taken in the Google Chrome throttle tool.

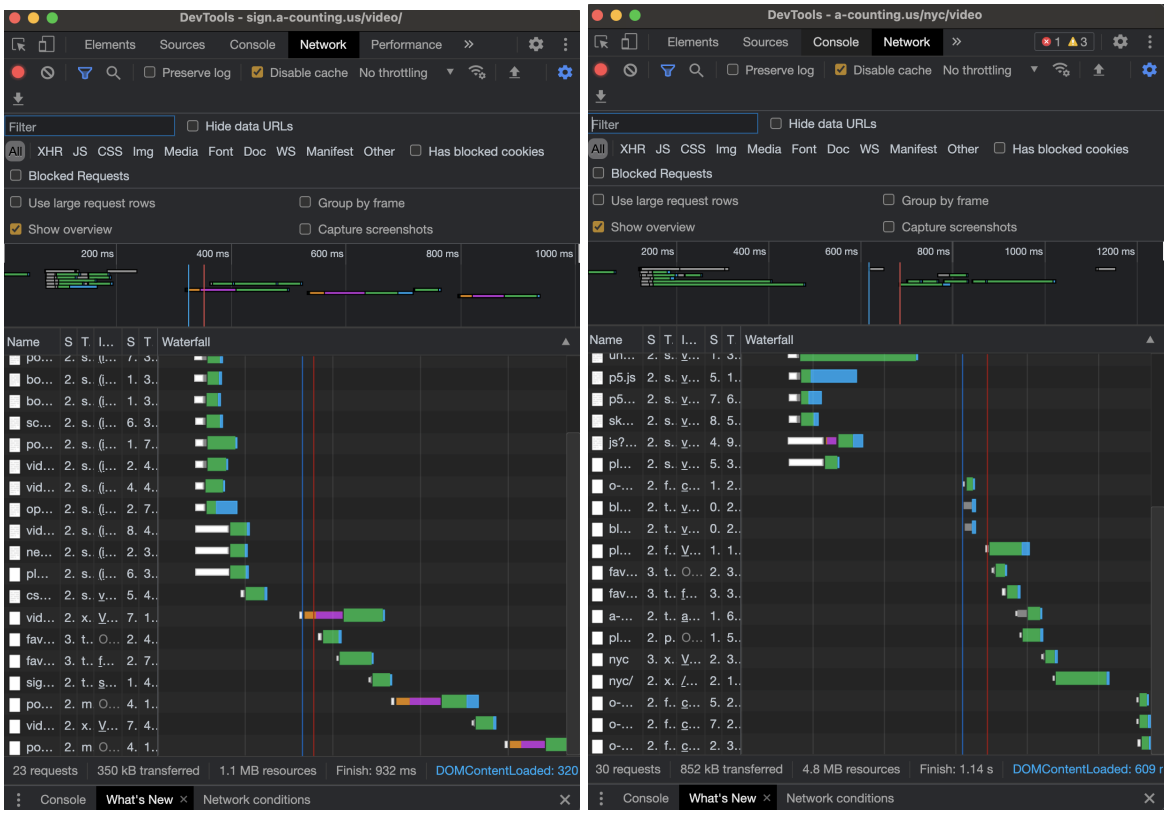


Figure: Load time and cascades from the Google Chrome network emulator, showing that sign (left) loads faster than the voice edition (blue line), as its media is cascaded differently.

4.11 Total system pipeline and novelty

For art does not transcend our day to day preoccupations; it brings us face to face with reality through the singularity of a relationship with the world, through a fiction.

Bourriaud -- Relational Aesthetics

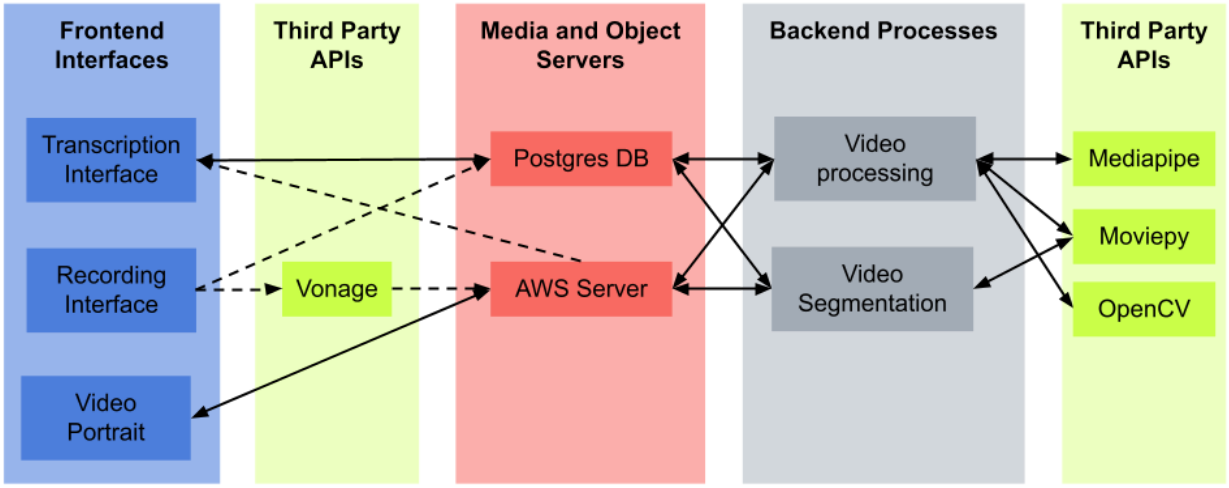


Figure: Higher level system diagram as seen in the beginning of section 4

There is much novelty and importance in making the above processes a cohesive pipeline with low cost and open source technologies. While no individual component of the above diagram is anything radically departed from existing literature, the use case of this project towards these technologies is unique. Furthermore, at time of writing, there does not seem to be existing methodologies for a variety of these components in this approachable manner.

The hope is that the availability of this thesis and its existing documentation is that these technical contributions can allow for more online artworks and explorations to expand what they do with video media.

5. Qualitative Contributions

Along with the technical contributions outlined above, this thesis also addresses and contributes to qualitative observations, methods, and lessons.

It is worth noting that a variety of email and social media outreach occurred for this project and others within the Poetic Justice group. Much of this is outside of the scope of this thesis document and is therefore not included, but please see the acknowledgements below regarding said outreach.

Special thanks to Tim Loh, PhD student MIT HASTS, for being an early tester and who provided signs 1-100 for the transcription interface testing for this user test. Additional special thanks to Max and Jack for additional early feedback.

Thanks to my UROP student, Skylar Kolisko for doing outreach and follow up coordination for the testing as well as quality assurance across multiple devices. Thanks to my advisor, Ekene and his administrative assistant Emily Pereria for making the gift cards possible for the participants of the user test to make for a better interface experience for users.

As well as to Cierra Martin, Aimee Burnett, and Charles Eppley for their development of participation guides and outreach materials for this and the vocal edition of the *A Counting* series.

5.1 User testing

A user test was conducted across 23 participants to determine the ease of use of the interfaces across a variety of individuals and devices.

This was a remote, fully digital, and asynchronous user test. Testers went to one link where they participated in 2 tests and had the option to record an additional video signing 1-10 and a testimony about their relationship with their language for demo and promotional purposes. The test took users between 12 - 25 minutes of their time and they were compensated with an Amazon gift card.

5.1.1: Test 1: Recording

This test had users record themselves signing eight numbers in any sign language with three different in between states with timed number prompts. The most pertinent question the user test was meant to address was the in-between state of the signs. Unlike spoken language, where one is either speaking or not -- signs need to be entered by hand. For our purposes, since participants would be signing 1-100 it was important to have a uniform in between states for putting together the numbers for the final portrait.

After recording, users answered a survey so we could collect data on the following parameters.

Recording interface parameters for user testing:

Parameter	Options for user feedback	Result from testing
Timing between number prompts	Too short Just right Too long	Just right
In between state between signs What in between state the users liked the most and liked the least (two separate questions)	Test 1: Making a fist Test 2: Moving hands outside the red box Test 3: Holding the signs	Holding the signs (pausing)

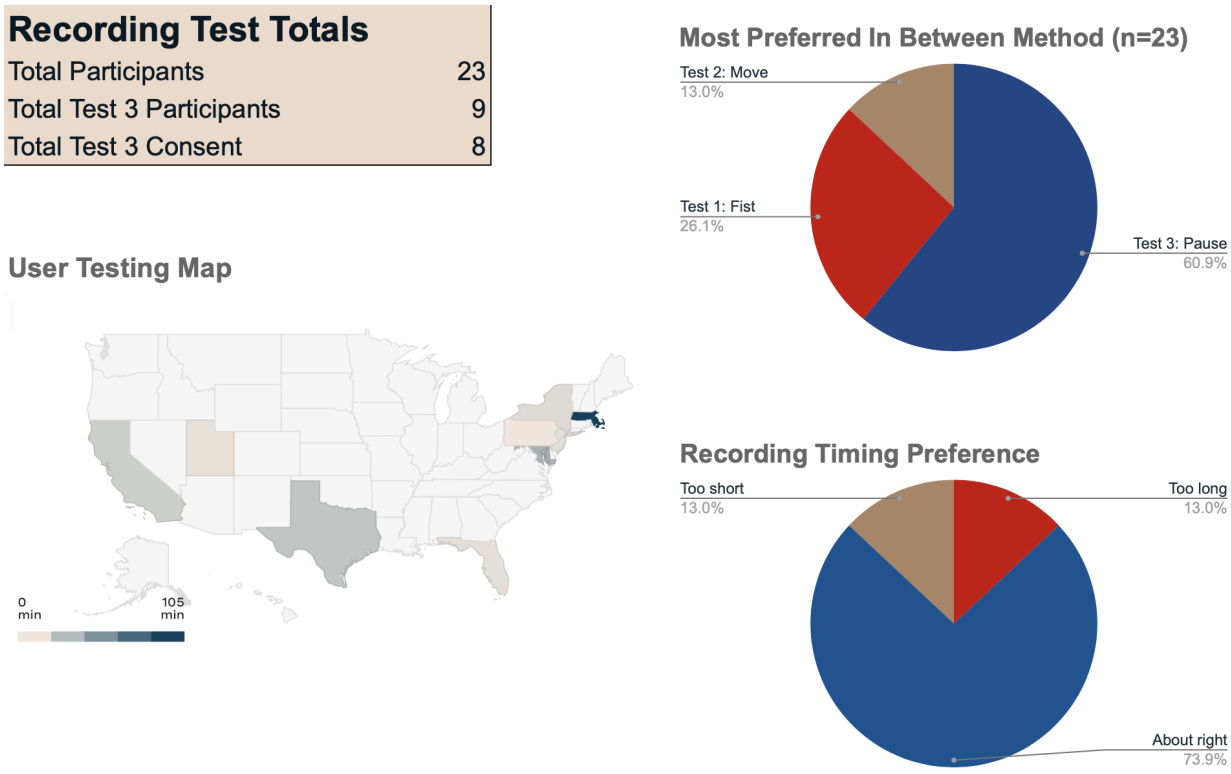


Figure: Graphical results from the Recording User Test, showing overall metrics and geographical distribution.

5.1.2: Test 2: Transcription

For this test, users transcribed a test count of 1-100 in ASL in the transcription interface. This test count was recorded by Timothy Loh on an early version of the system.

After transcribing 1-100, users answered another survey regarding the following parameters.

Table: Transcribing interface parameters for user testing

Parameter	Options for user feedback	Result from testing
Perceived ease of use: “How easy was it to transcribe?”	1. Very easy 2. Easy 3. Somewhat easy 4. Somewhat difficult 5. Difficult 6. Very difficult	Somewhat easy
Perceived accuracy of the sign slicing: “How was the clipping of each video? For example, was the start or end of the sign cut out?”	1. Mostly accurate 2. Accurate 3. Somewhat accurate 4. Somewhat inaccurate 5. Inaccurate 6. Mostly inaccurate	Somewhat accurate

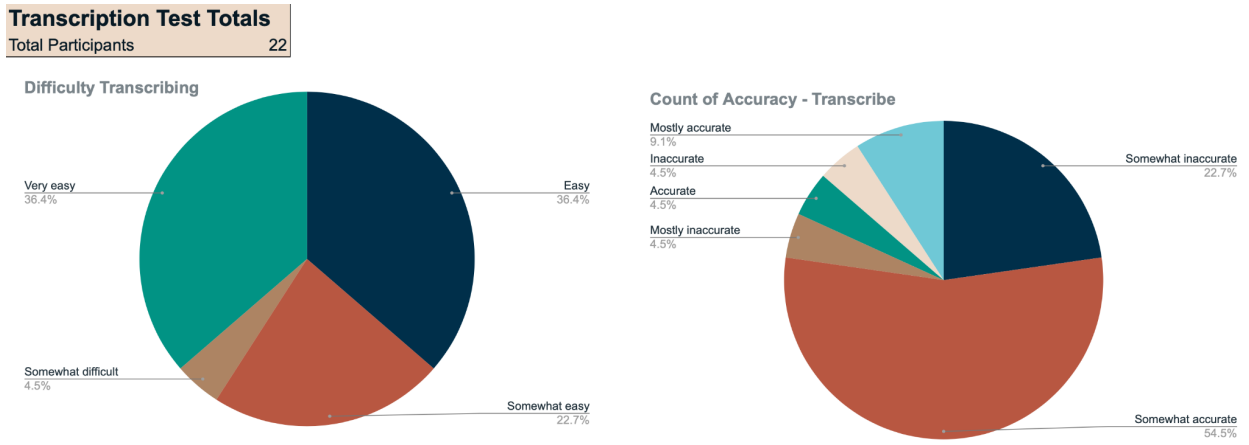


Figure: Graphical output showing the summary of the transcription user test.

The third test, an option test, allowed users to record themselves signing 1-10 in any sign language

Both exit surveys use forced choice scale methods, which does not allow for a neutral or unbiased result. This allows us to truly understand which side of the decision line the user fell for each parameter.

Of the 23 users, 7 left additional written feedback about the interfaces. This is summarized below (single word feedback or notes about the participant's recordings i.e. if they messed up on a number are taken out)

Feedback: Transcription	**Feedback: Recording**
I liked it! It didn't take me too long and I did it on a computer. There were some numbers I was on the fence with, so it might be good to let people know what to do in that situation. I wish I could enter text too instead of just dropdowns for invalid videos. Also it might be good to have in the instructions that an algorithm did this, in case someone didn't read the home page in depth. (transcription)	Thank you again! | This is a really cool project! The interface was easy to use and worked well but I think it will be a lot better to test for pauses because doing other things for a lot more signs would be too tiring.
Thank you! This is a really cool project! I understand you need to have a marker to differentiate signs. In American Sign Language, making a fist between signs/numbers is a little unnatural, so if it's not too difficult, perhaps you could instead recognize the signers hand going limp, as well as pauses (when the end of a number is held before moving on to the next number). I think this would help since this is what I usually see ASL signers do (like in this video that I helped transcribe).	I missed whether you wanted repetitive signs or sign only once (recording)
I felt there was some "clipping" at the beginning and the ending of the video. There was not enough time for the sign to be 100% clear. (transcribe)	This is a great innovation anyways. It would love sign language users to be more receptive in their skills.
Some videos are not cut up at the right pace of sign	

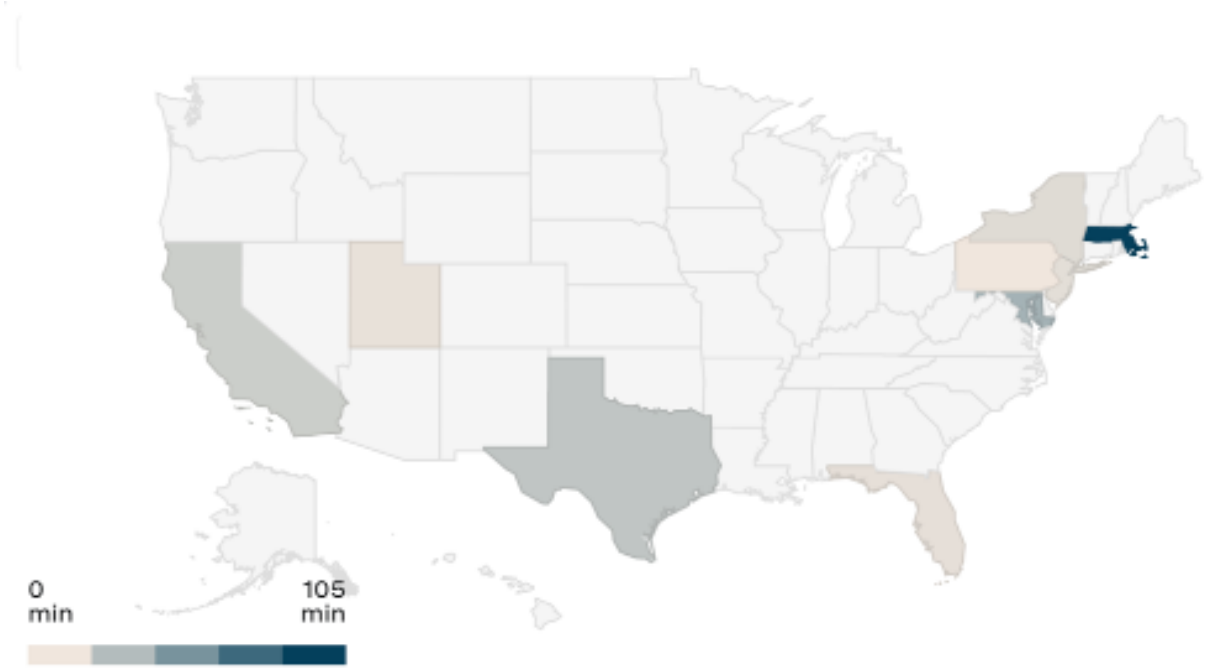
9 participants took part in Test 3 (signing 1-10). 7 of these 9 left a testimony. These recordings were used for testing and demonstration purposes. Below are these testimonies transcribed in written English of what users for this test said about sign language:

Hello, my name is Janice. My sign name is *signs name.* I love sign language. I have a lot of deaf friends. I also offer my interpreting services for the church. Goodbye.
Hello. I love sign language because it is a great mode of communication that we use. Unlike speaking, sign language allows us to express ourselves through body language such as facial expressions and shoulder movements, essentially it is a 3D language. It is just really wonderful.
Hello. ASL is my mode of communication because I am deaf. I really must have ASL in order to communicate. Goodbye.
I use sign language with my partner. In other situations, I use my voice and listening skills pretty good. So, for me I am currently living in my immediate circle and surroundings. Sign language is love to me. It allows me to communicate with my partner. Also, I have deaf friends, interpreters and KODAs. Also, sign language is fun, accessible, exciting, collaborative and mutual. It is nice to be a part of the community. The experience is rich, with

background, experience and people who persevered through hardships and want to be here. Sign language allows us to be here.
ASL is my first language. I really love ASL. I am fluent as it is my language. Sending love.
So, sign language is a part of my culture, community and it helps me to communicate. Sign language is a big part of my life.
Sign language makes me who I am. That is that. Thank you.

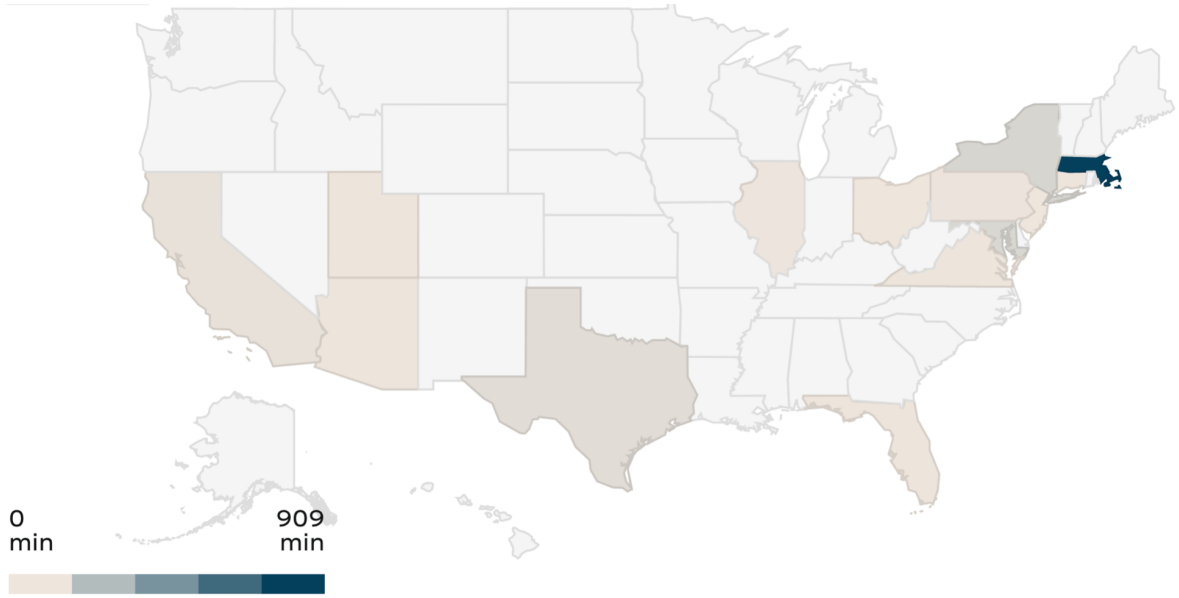
As many theses this year will recount, it was difficult to perform a user test during a pandemic. It is impossible to control for so many variables and it was difficult to gather participants during such a saturated digital time and with no in-person interaction. All data within this thesis should be read with this context in mind.

However, one large positive of having a fully digital remote and asynchronous user test is that participants that would not have been able to participate in person (whether it be due to timing or location) can participate. For this user test, we had participation from 9 different states, as one can see from our Vonage usage geomap:



This wouldn't have been possible in a traditional user test and represents another layer of linguistic diversity for this project.

At the time of writing this (June 2021), the current distribution of the project geographically is reflected below, showing the work has participants from 14 states now!



Overall, given the summaries above, it can be concluded that users were satisfied with the interfaces in context of usability and legibility. There were no users that considered the system unusable. Furthermore, a clear preference for an in-between state and timing was made.

5.2 Educational contributions

This section summarizes a workshop I developed and taught at Gallaudet University in collaboration with Professor Max Kazemzadeh. Special thanks to Max and his students for having me. Thanks to Zach Lieberman for feedback and for introducing me to Max. Thanks to Joyful Signing for such professional interpretation services for office hours. All closed captioning was done by the service Rev. Thank you Skylar Kolisko for developing the glossaries for the lectures, trying out the lessons and templates, and making flyers for office hours.

Materials are available here: https://ninalutz.github.io/gallaudet-creative-coding/ but the section below summarizes the programming developed and reflections from the overall experience.

5.2.1: Context of Gallaudet

Gallaudet was founded in 1865 and is the only liberal arts institution of its size dedicated to the signing community [gallaudet.edu]. It is featured in a variety of documentaries, such as “Deaf U” on Netflix, and is a leading scholarly authority on Deaf culture and American Sign Language [Netflix]. Its mission is to provide a quality liberal arts education through American Sign Language:

> *Gallaudet University, federally chartered in 1864, is a bilingual, diverse, multicultural institution of higher education that ensures the intellectual and professional advancement of deaf and hard of hearing individuals through American Sign Language and English. - Mission statement, Gallaudet.edu*

For many students, Gallaudet is the first time they are in an environment where ASL is the dominant language and they are interacting with a majority Deaf population. Undergraduates come from all over the country, and even the world, to receive an education within Gallaudet.

Gallaudet offers a variety of Bachelors, Masters, and Doctoral programs. For the purposes of this thesis, I lectured in undergraduate programming within the BA in Art and Media Design program for undergraduate students.

5.2.2: Planning materials

I had several meetings with Max in order to best understand where a guest module would best work within his programming for his multiple courses.

There was much need for planning for our collaboration. The goal was to have me guest lecture to introduce students to creative coding concepts that they were covering in Max's class, alongside computer vision and the technical breadth of *A Counting: Sign Language.* This was also within the context of wanting to foster future collaboration between Gallaudet and the MIT Media Lab, particularly within creative coding and Zach Lieberman's Future Sketches group.

I gave Max different options for modules and eventually we settled on doing a guest module in his Web Design 2 class. This would be two 3 hour lectures and 2 optional office hour sessions where students could come to ask for help. Max's Web Design 2 class utilized p5.js to introduce students to the fundamentals of web programming through a creative coding lens. My particular module would introduce object oriented programming with APIs and discuss the computer vision and research methodologies utilized in *A Counting: Sign Language.*

Furthermore, we decided on me doing a guest lecture regarding creative coding, computer vision, and *A Counting* in his Advanced Digital Media course, a course that examines the higher level trends in digital media with independent art projects and serves as a general studies elective along with being an elective for the Media Arts students.

Along with these live sessions and office hours, I also prepared templates for homework for the students along with close caption video walkthroughs of the code.

5.2.3: Guest module and lecture

The first lecture of the guest module covered sketching with code and the importance of iteration in coding and sketching.

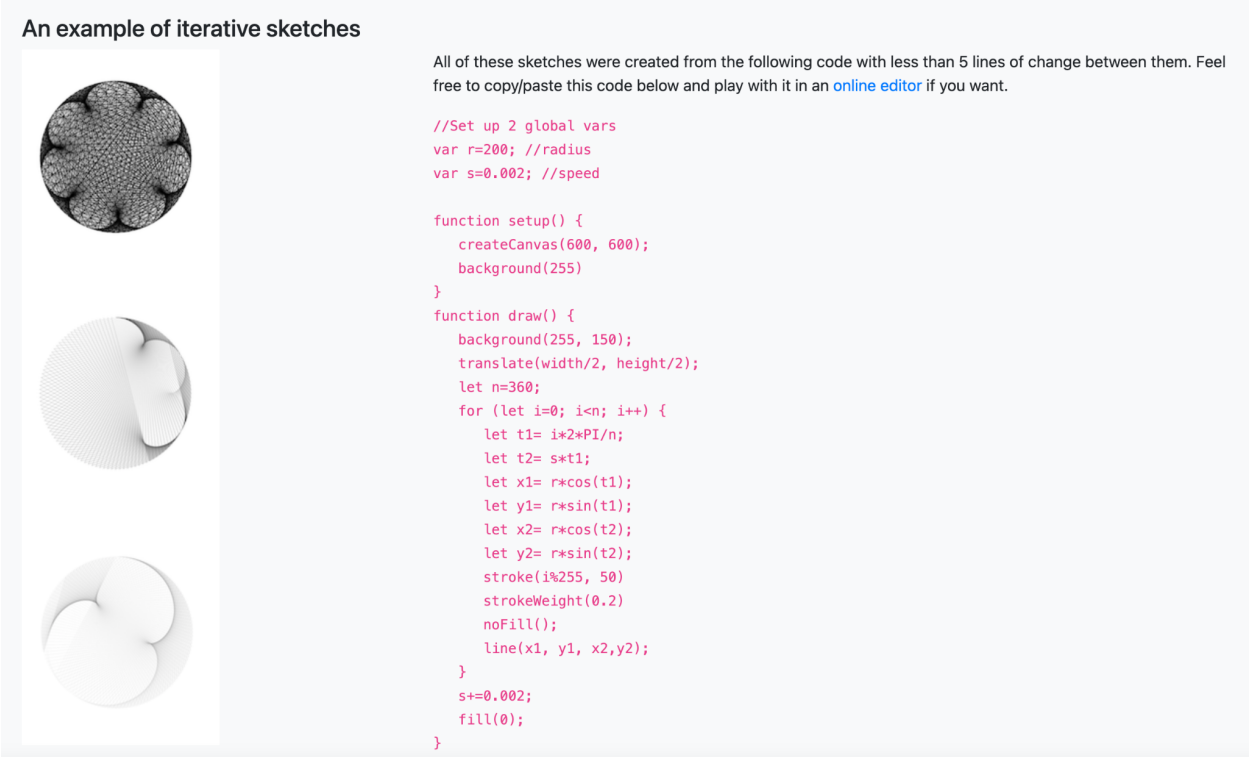


Figure: Excerpt from module website

I then segwayed into the idea of objects in code making it easier to do iteration and to make sketches that interact with data on the internet. This is where I introduced the concept of APIs and how servers can transmit data from APIs to p5.js sketches.

I have taught this type of content before in p5.js, and I have done it with weather sketches as there are several simple weather APIs to connect to and students are able to understand and generally enjoy weather data in their sketches.

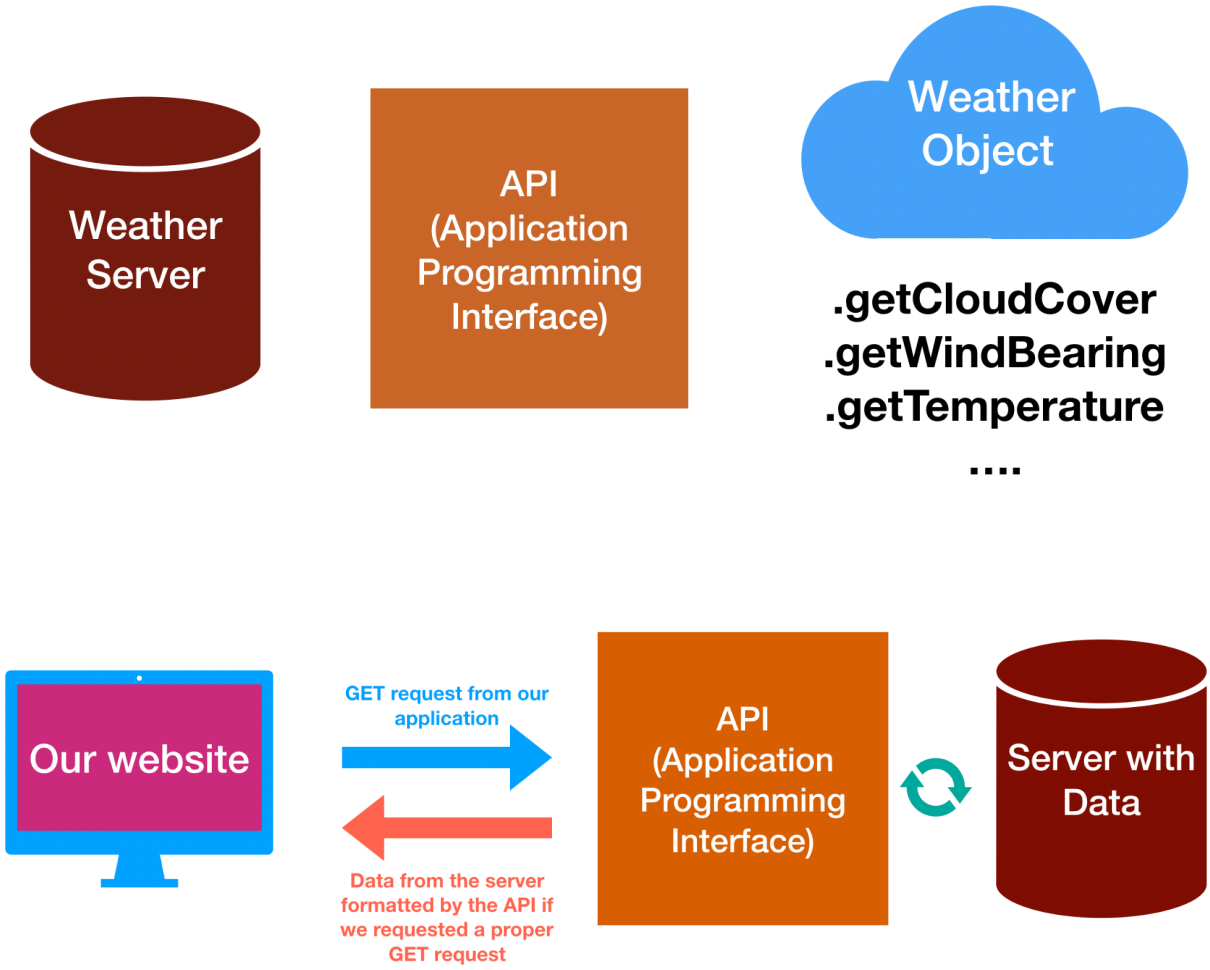


Figure: Graphics from slides and website to explain APIs and servers on the web for the sake of the lecture and website documentation

Overall, I did walkthroughs of the code in class with a variety of weather sketches that I made for examples and for homework templates for students to make sketches based on any location that they wanted.

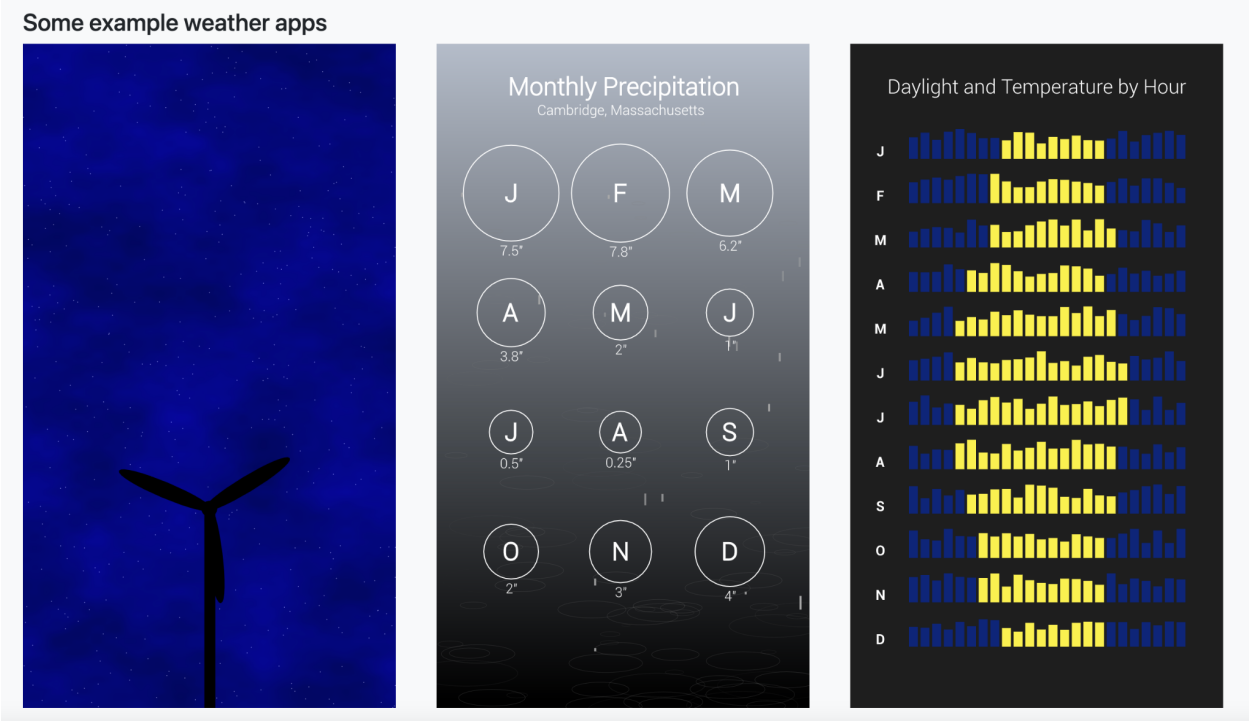


Figure: Example weather sketches I made templates and walkthrough videos for

In my next lecture, I covered more information about the objects we defined and gave a higher level overview of how the Vonage API works in *A Counting Sign Language.*

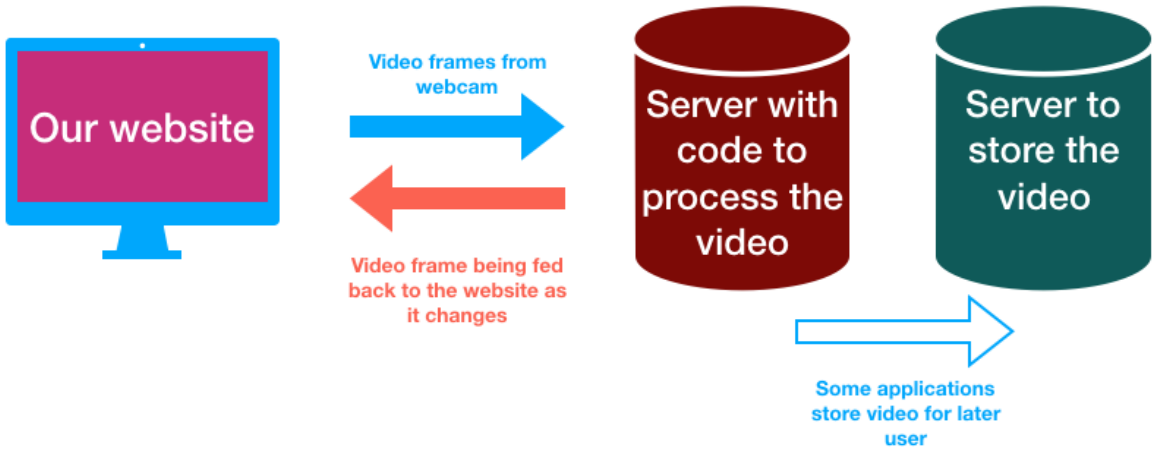


Figure: API diagram for video streaming across the web

I also gave an example of the Mediapipe API for artistic purposes in p5.js. Students greatly enjoyed this example overall:

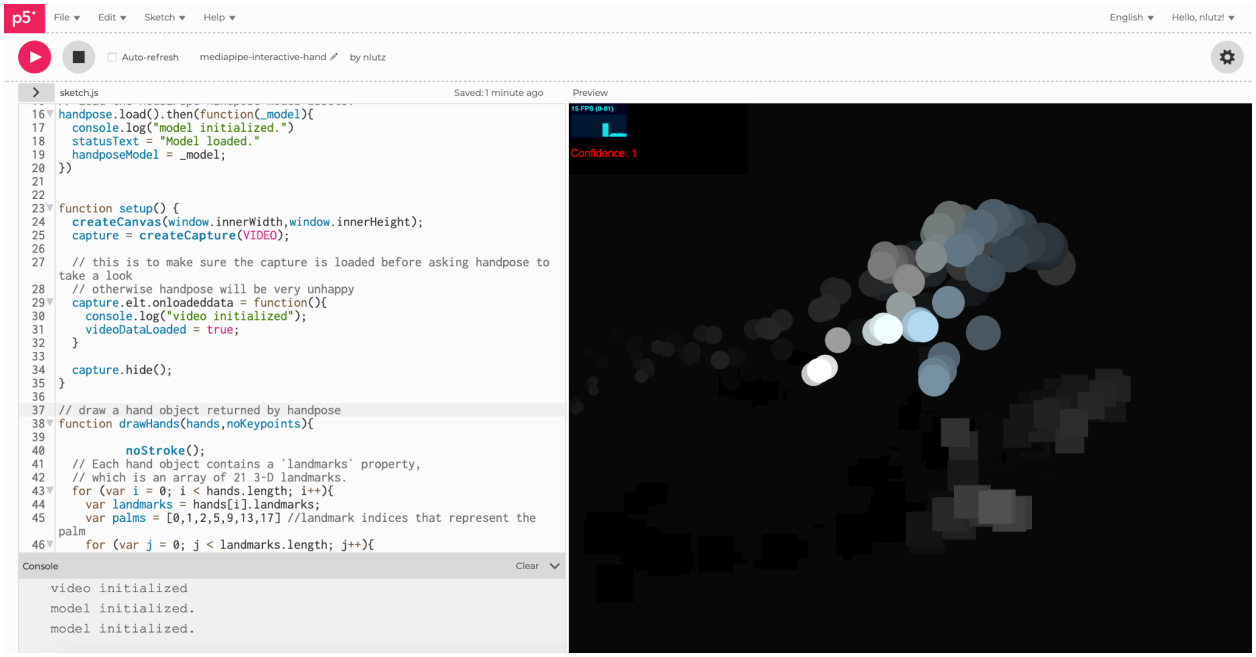


Figure: Screenshot of Mediapipe and p5.js example, which students covered in class

I then described in more detail the pipeline and mission of *A Counting: Sign Language* and gave students a chance to try it during their break.

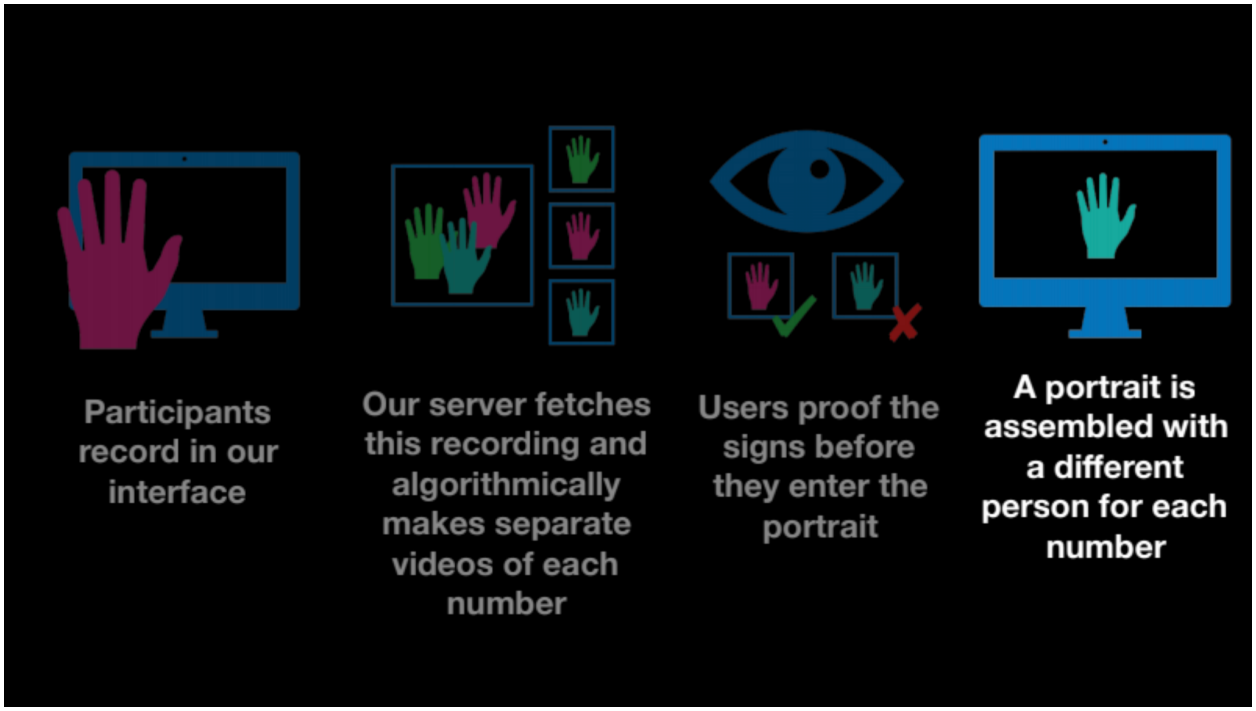


Figure: One of the diagrams about A Counting: Sign Language shown in lecture

I also held office hours for this module so students could come and ask questions. Some students did come by, which was great. ASL Interpretation for these office hours was provided by Joyful Signing.

Along with teaching this guest module, I also have the pleasure of giving a guest lecture in Advanced Digital Media. The guest lecture covered a variety of topics within creative coding, projection mapping, user research, and *A Counting: Sign Language.*

I tried to make this lecture highly visual and bring in a variety of videos with closed captioning so students could focus on one video stream and not on the interpreter and slides at once.

Overall it was an extremely educational experience and I feel exceptionally grateful to have been invited to teach at Gallaudet for the context of this thesis.

5.2.4: Conclusions

Teaching during a pandemic is difficult. Learning during a pandemic? Even more difficult.

Online instruction for many universities this academic year was exceedingly difficult for instructors and students to adapt to during COVID-19. Student mental health plummeted as many students went to more distracting and less supportive homes as they adapted to digital learning in an increasingly racking and uncertain world. Many students fell behind on their degrees, struggling to maintain their motivation and current trajectories in these changing environments and uncertain times.

Gallaudet University was no exception, and if anything, many of their students and instructors faced increased burdens. For many Gallaudet students, they face language deprivation at home -- families that will not or cannot sign with them. Other students face additional challenges receiving the interpretation that they may need in order to do digital learning.

And throughout this age of Zoom university, many of us have heard stories of accessibility. Indeed, many events and conferences have been made more accessible than ever before -- for individuals across the world and with increasing visibility and use of accessibility options such as captioning and interpretation and asynchronous programming.

Unfortunately though, my biggest take away from this experience is that for Deaf and students that use sign language, Zoom university is often less accessible and more difficult. These students often have one screen (often just their laptop or a tablet) and are trying to watch and pay attention to: lecture slides, someone signing, and their own work (if relevant). This is extremely tiring and overwhelming and results in many breaks during lectures.

This means that students are even more likely to fall behind as much of the work cannot be done during class time and they must do it later -- in a variety of different households and learning environments. For many students at Gallaudet, English may not be their first language.

This makes doing this work outside of class and trying to keep up with all of it even more difficult for many students.

Furthermore, for students on Zoom, there are many features that are just less inclusive than one may think. For example, Zoom's greenboxes, which move to the person talking, do not have any visual input and only travel to sound, no movement. There also is not a way in Zoom to control which user is "spotlighted" if you are not the host, so it can be difficult to keep track of who is signing when there are also slides being shared on the video call. Especially when you are looking for smaller hand movements in tiled grids of videos, on a variety of screen sizes and resolutions.

All of this and more leads to exhaustion and demotivation and was perhaps the most apparent reason why we had such little participation in *A Counting: Sign Language.* I structured the lecture to add this call to participation within a 15 minute break (as the project only takes 5 minutes). Students were, relatably, more interested in a break than participating in any art project. Additionally, it was difficult to explain the context of an art project and participating in it while they were also trying to understand the technology behind it in the context of their work.

While participation in the *A Counting* artwork from these units was incredibly low, I greatly enjoyed the process of preparing, teaching, and hosting office hours with Max and his students. This engagement greatly informed a lot of the understanding regarding signed languages and accessibility for the qualitative critique and positioning of much of this thesis. Furthermore, it is my hope that this may lead to future interfacing between MIT Media Lab and Gallaudet University, as many of Max's students were very interested in this type of work, some of them sending me various project ideas they would love to develop further.

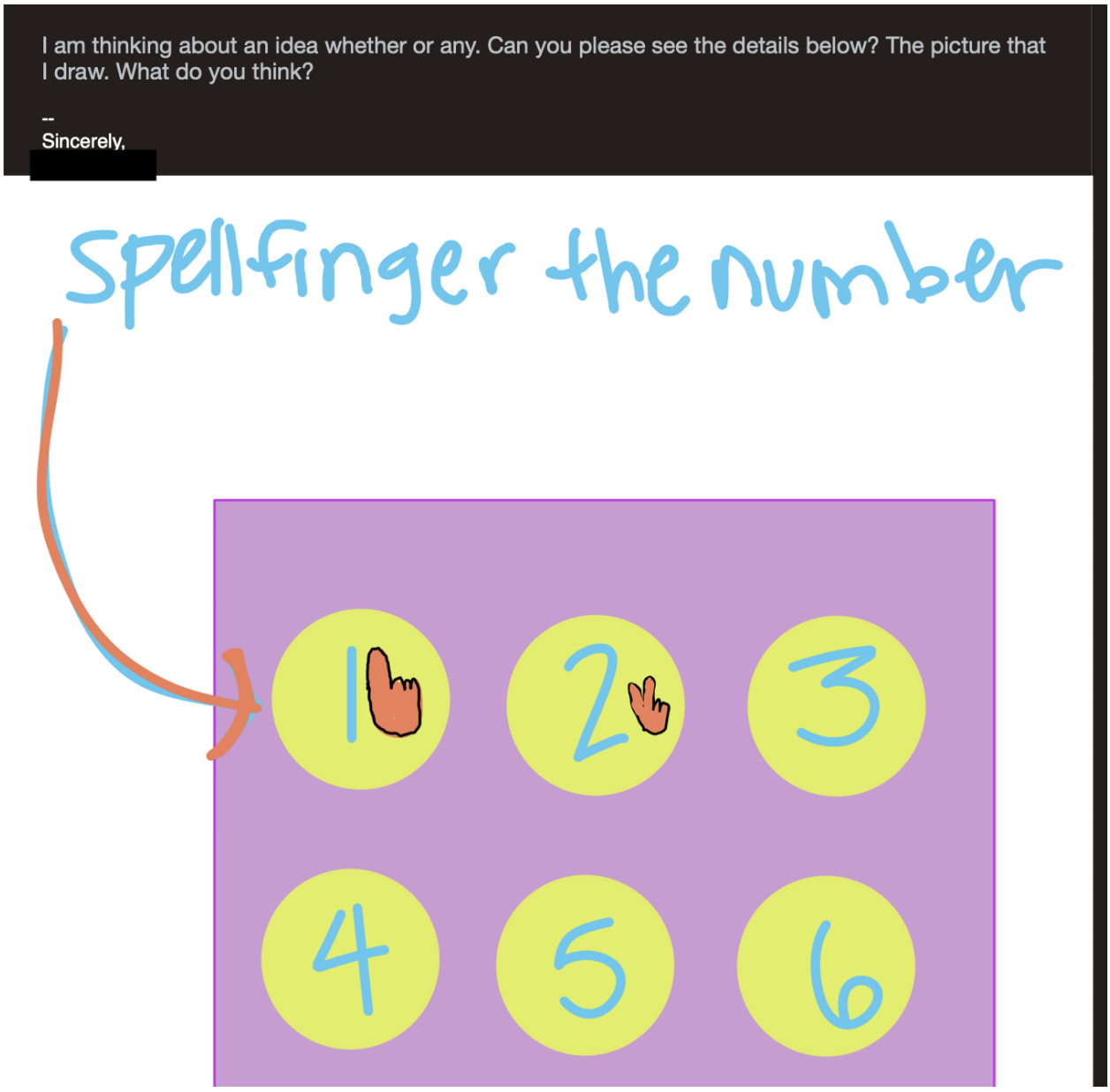


Figure: Student email from Gallaudet after my creative coding lecture, inspired directly from A Counting. The student wanted to build something like this in p5.js as part of the course. Student engagement like this is what all of us educators always want to see.

5.3 Video portrait design

Overall, the goal for this video portrait was to achieve unity amongst the separate clips while also preserving the authenticity of recordings. There was never any intention of abstracting away from the original footage itself for the final portrait.

The priorities to achieve unity amongst the clips were to center the hands and have all faces removed in order to have the focus be on different hands in one unified count.

Any processing was designed with a very light touch, as outlined in the color balancing section in Section 4. This was to maintain the goal of authenticity.

The aspect ratios considered were all meant to be landscape ratios, for exhibition purposes and for keeping consistency with the voice edition.

Figure: Showing a variety of test samples in different aspect ratios, from 4:3, 3:2, 16:10, 16:9 in order

However, since *A Counting* is a public participatory artwork, the video portrait is available on the web. This means that videos also had to be generated for portrait orientation (16:9) devices. This was another layer of video design that influenced the design and implementation of these video portraits.

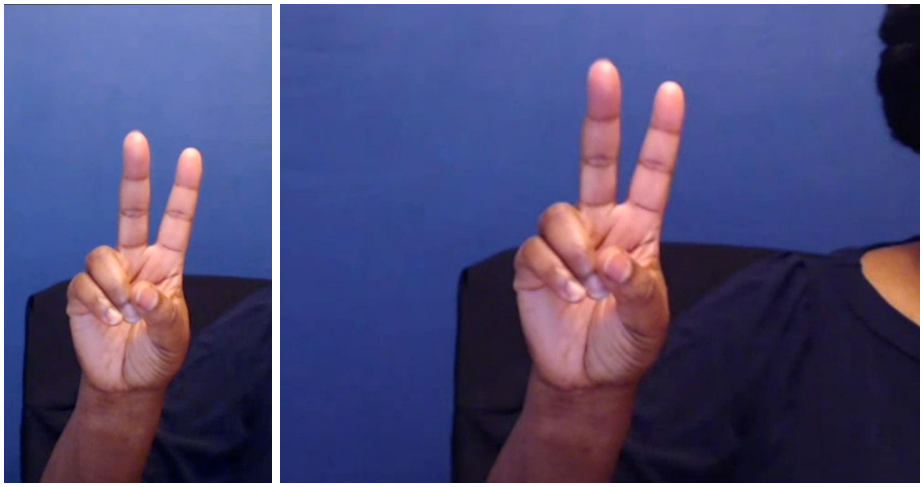


Figure: Example of a number in both its portrait (left) and landscape format (right)

Overall, these design decisions were made to mimic the voice edition and also in respect to what would look best in a gallery setting and how typical video installations are displayed (i.e. a landscape video being more eye-catching than a smaller square video).

6. Analysis and Evaluation

I am extremely grateful for Sheena Stuart-Milburn's feedback of the project, visual identity, and website as her unique intersectionality of being a graphic designer and Deaf individual added perspective that was invaluable.

Other individuals: Max, Timothy, Gregor, Jack, and more were invaluable with their perspectives on the framing and analysis of this work. Thanks to Skylar Kolisko, my UROP student, who set up Google Analytics for the website.

There are some quantitative ways to evaluate components of this thesis, particularly the web application component.

6.1: Website analytics and usage patterns

At the time of writing, the system has over 100 unique participants, meaning that over 100 people have started recording or transcribing videos. Furthermore, there are 19 full counts to 100 in American Sign Language, and 9 additional counts to 10 from the optional test offered to users. Of course, more individuals and bots go to the website as a whole.

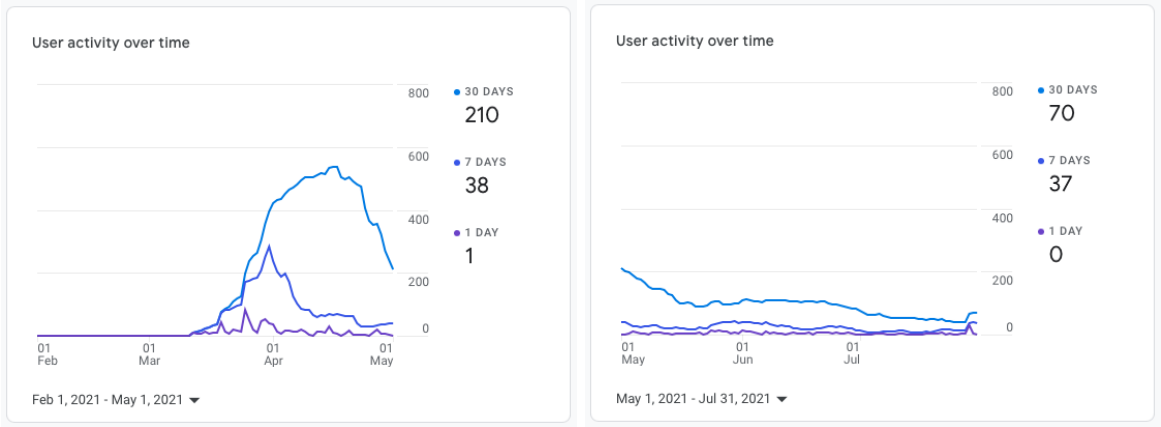


Figure: Screenshot showing heavy traffic earlier in the summer after release in April and sustained traffic of consistent 100-200+ engaged users a day. Source: Google Analytics

6.2: Software performance and results

There are many quantitative measures by which software can be evaluated. For the scope of this thesis, particularly as it is not an engineering thesis, performance (speed and memory) of the web application that enables the artwork are considered. The algorithmic efficiency of the video segmentation and processing approaches have been covered in previous sections.

6.2.1: Recording and transcription interface performance

The recording interface guides participants through recording themselves signing 1-100 in any sign language. It takes a little over 5 minutes, longer if a participant chooses to leave a longer optional testimony at the end.

The load time for the page is under 5 seconds on all networks in the Google Chrome emulator, overall quite efficient. Because this is a server-side implementation of recording, despite over 5MB of total information being transferred during a recording session, this does not occur or affect the client device and is not of concern.

The transcription page is similar, although slightly slower on some networks, hovering just above 5 seconds on 3G speeds, and thus having a loader on the interface design to accommodate this. Over 5MB of data is transferred for this interface as well.

6.2.3: Web application architecture metrics and performance

The web application container of a Heroku application consists of the code, libraries, caches, and binaries to run the application. This controls how responsive new builds are to the application as well as gives an indication of how flexible the codebase is to future developments. This is called a slug in Heroku and the platform has a hard limit of 500 MB slugs for compiling, but recommends staying under 300 MB [Heroku].

The current slug size for *A Counting: Sign Language* at time of writing is 263.5 MB, leaving plenty of room for future improvements and developments.

6.3: Accessibility

It is important that the website is as accessible as possible for as many people as possible. There are a variety of web accessibility standards and tools for seeing how accessible a website may be for screen readers and other accessibility aids [W3.org]. One such tool is Wave which any website can be entered into in order to see any issues on the page that may be present:

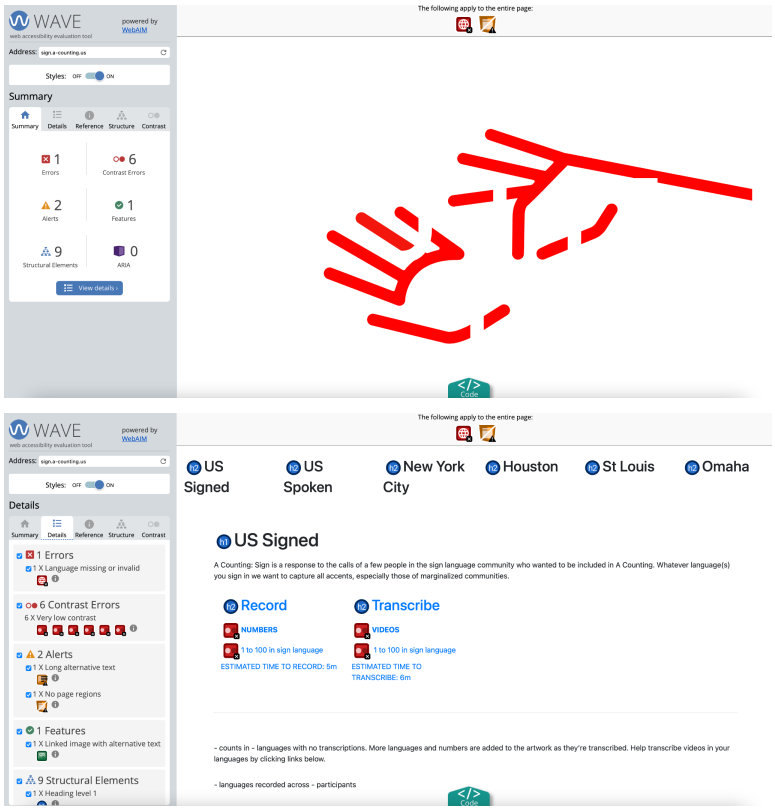


Figure: A Counting: Sign Language in a WAVE (web accessibility evaluation tool) preview showing various accessibility details, errors, and suggested changes.

Accessibility for this website in particular is quite important since a variety of Deaf individuals and other individuals who use signed language use screen readers and other assistive devices. Some important principles for this project's website include but are not limited to the table below:

Table: Web accessibility features in *A Counting: Sign Language*

Feature	User needs and experience	Implementation status
High contrast colors	Many users need high contrast colors for their vision in order to navigate visual information	This is implemented throughout the *A Counting* visual identity and in the recording interface
Alt text and captions on any media	Users that utilize screen readers or views with screen information like text rely on alt text for information about multimedia content on websites	Alt text is provided for logo animations on the landing page of the website.
Ability to zoom in without losing information	Users need to be able to have access to all the information while being able to view it in a large font for vision purposes.	This is handed by the containers set in place by Bootstrap's grid structure
Timing of dynamic content	Users have enough time to read and interact with any dynamic content and instructions on the website	This is done by having buttons for the user to press when they are ready to start recording

Qualitatively, the main accessibility barrier at time of writing is how this project and its participation roles are communicated. From instructions on the website to calls to action to advertising -- this is within itself a part of accessibility. While the interface itself had good responses during user tests regarding usability, many potential participants are often confused by the calls to action for this work:

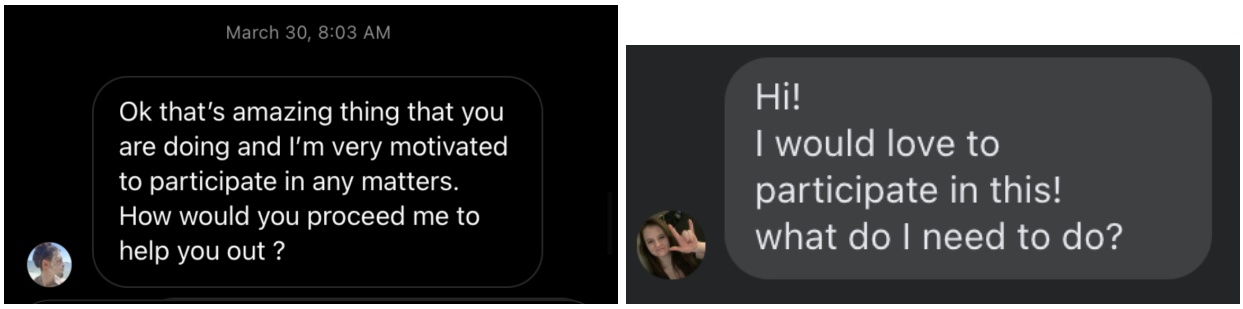


Figure: Screenshots of people being confused about what to do with the interface based off the call to action sent to them via social media messages.

This is partially due to the fact that for many of these participants, English is not their first language. Furthermore, the linguistic association for most people with words like “portrait” are not of language but with faces. Words like “transcribe” often have an audio connotation and may also need to be translated instructions in ASL for the website for it to be fully bilingual and accessible.

Additionally, it is worth mentioning that none of the user testing participants utilized screen readers and all were recruited through written English messaging on Instagram and email. This in itself means that there are and will be more accessibility improvements that emerge.

6.4: Ethical considerations of technologies and methods used

To avoid the trivialization of ethical concerns of technology development, implementation, and potential for exploitation, such matters have been interwoven throughout this document. This section is to serve as a summary as well as a critical ethical placement of this work and its methods.

The first focal point of these considerations lie within the technologies that enable this artwork and considering the ethical implications of their use but also the platform and organization behind them. This is summarized below:

Table: Summary of ethical concerns and measures for technologies utilized in *A Counting: Sign Language*

Technology	**Ethical Concerns**	**Solutions within this thesis**
OpenCV	OpenCV is a powerful software that has the potential and is currently being used for malicious computer vision tools.	This thesis does not support or contribute to any public OpenCV projects nor does it contribute data to these efforts.
Information collection and usage	Information is collected: names, zip codes, contact information, metadata via video, etc	Security measures and not exposing this data, while also adding a GDPR compliant form acknowledgement in the interface itself [GDPR]

There is additionally, the carbon footprint of web technologies such as servers used in this project. While this application is not being served on a green hosting platform with carbon offsets, its implementation and scale do not suggest a large footprint. Additionally, according to a web carbon calculator [WebsiteCarbon], *A Counting: Sign Language* is greener than 85% of the web.

With regards to the ethical groundings of the work itself, the thesis speaks to this at length on how this work developed from precedents and considerations and requests of relevant frameworks and theories.

The work itself is situated in a careful space to not push any agenda or discrete value judgments on signed languages and those who utilize them and to prevent potential harm wherever possible with precautions of framing and security.

6.5: User responses

In a participatory artwork such as this, the most important critique is that of the participants' experiences in its creation.

Language is an intimate and integral part of many people’s identities. In both the voice and signed edition, we received several testimonies regarding this. And this participant response, criticism, and engagement are some of the most valuable critiques for this type of work.

Figure 2: Screenshots of video testimonials from *A Counting* vocal edition

When this project first started and I was seeking early user testers, I was at first trying to reach out to the Harvard and MIT signing clubs. This is when I discovered much of the complexities around signed languages identity and how this might interface with this work. For example, it was pointed out very early on that non-native signers would not want to contribute recordings for the portraits:

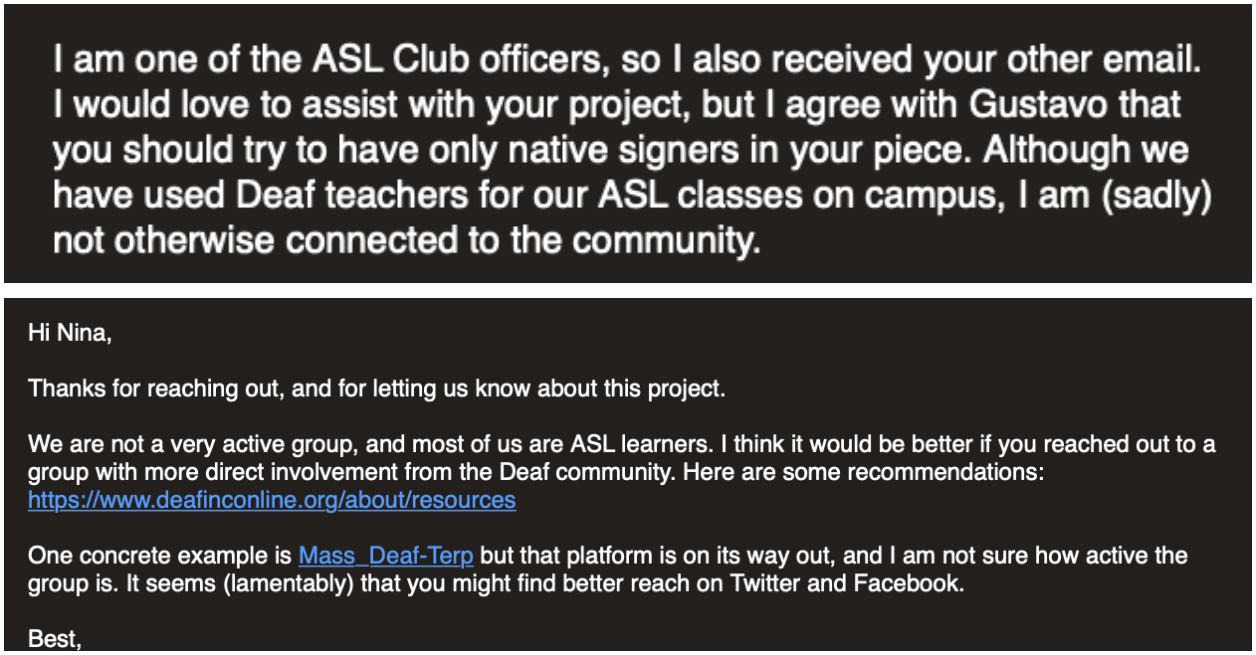


Figure: Two officers from the MIT ASL replying to an email regarding being non native signers and not wanting to partake. These early emails were very helpful for framing and future outreach strategies for this project.

This is perhaps why one of the first transcriptions received was from "a hearing signer", an individual that wanted to make this clear to us.

While the user test provided insight into opinions of the project and of users' relationships with sign language, many of which were very powerful and beautiful statements about sign language and identity, much can be gained from seeing the responses to the artwork from various outreach and calls to participation.

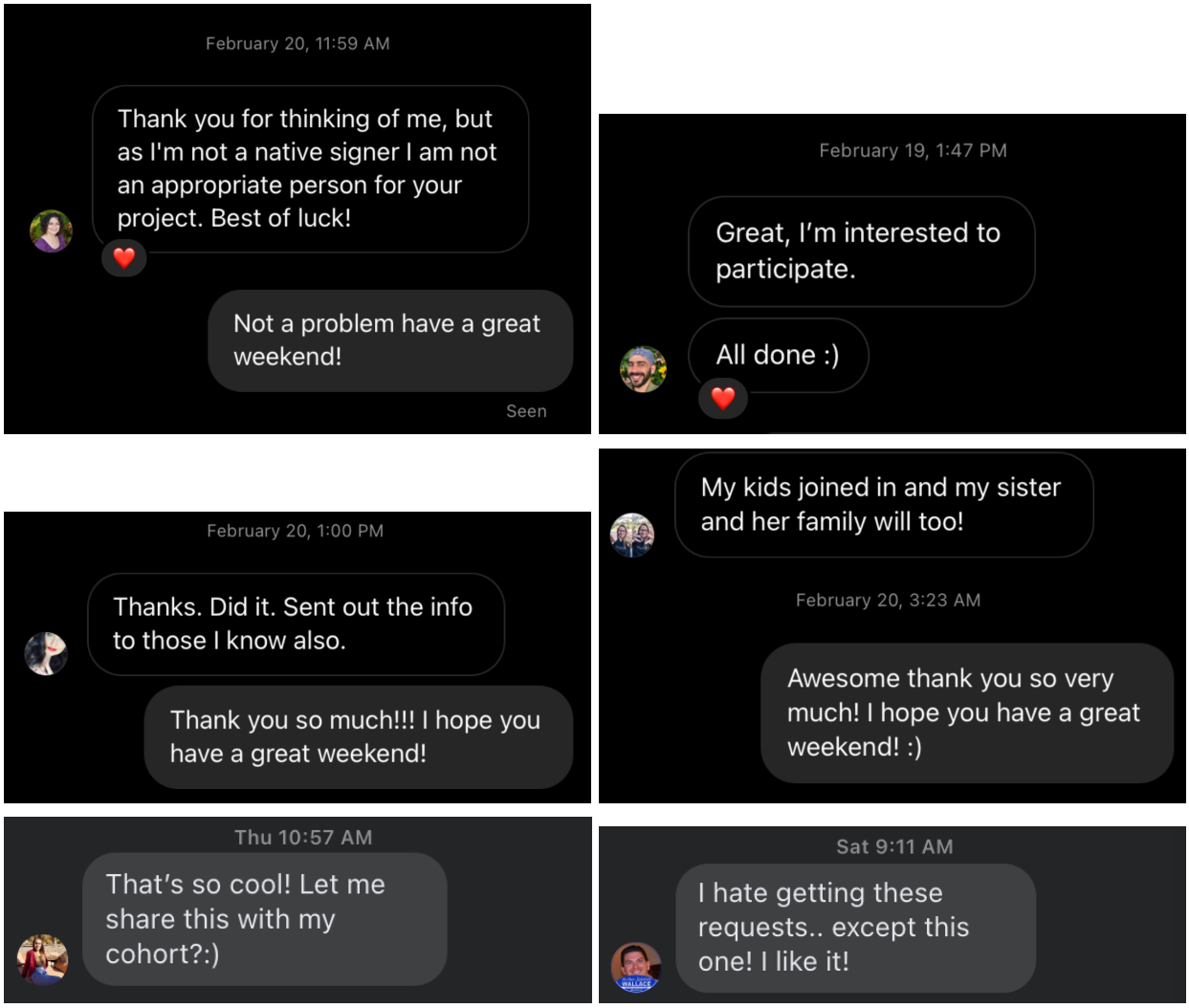


Figure: Collection of responses. Many responses overall were positive or with individuals who did not feel they were appropriate subjects. Additionally, many individuals were confused about how to participate but were excited about the project itself.

Along with the distinction, there were those who had questions regarding the involvement of Deaf individuals with the project. The response to this was to regard the on the ground interviews and user testing completed, as well as framing the project as an option for sign language to be present in an artwork that did not originally accommodate it, rather than a statement or suggestion about sign language.

Overally, responses from individuals have been fairly positive, but only time and further participation will tell how the reception and framing of this and other participatory works may change.

6.6: Improvements and future work

This work has much space for improvement and opens new doors and possibilities for potential future works.

6.6.1: Segmenting of individual signs

The following modifications could be made at a later date regarding the slicing errors discussed in 4.5.3, while keeping the methodology of slicing by the timing:

1. Modifying the clips by cutting margins on either end when slicing, most likely by erring on the side of cutting from the end of the clip more than the beginning since there is always going to be an entrance to the sign
2. Modifying the instructions of the transcription interface such that users can select multiple signs (e.g: 11, 11 and 12, etc).

The segmentation of the individual signs does not have to rely on the timing of the interface. It could be done by motion detection. As explained in section 4.5, this is easy to do when Mediapipe detects a hand, but is non negligible for when a hand is not detected. This solution was out of the scope of this thesis, but some experimentation was done towards it.

Firstly, it is difficult to predict when and why Mediapipe may go wrong. After all, it works very well on a variety of video resolutions, lighting conditions, and skin tones:

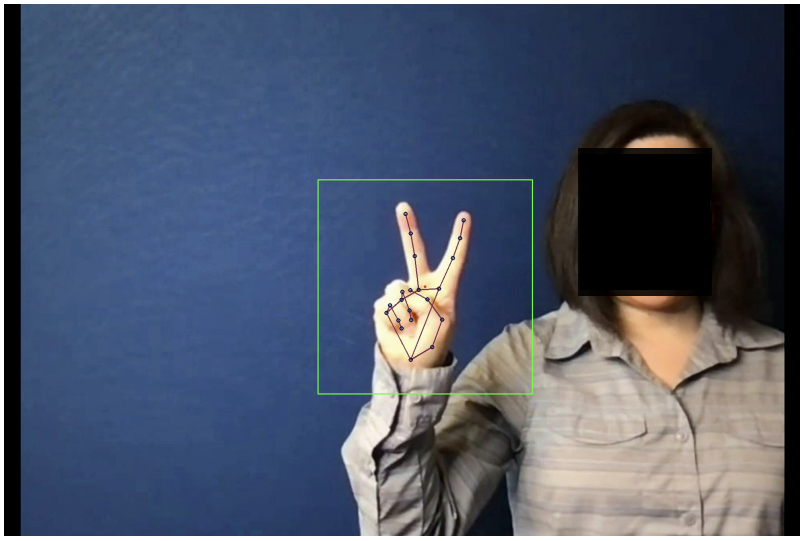


Figure: Examples of samples that Mediapipe successfully detects and finds hand landmarks

Unfortunately, Mediapipe does have issues on frames that should be easy to detect, as they are clear and well lit:

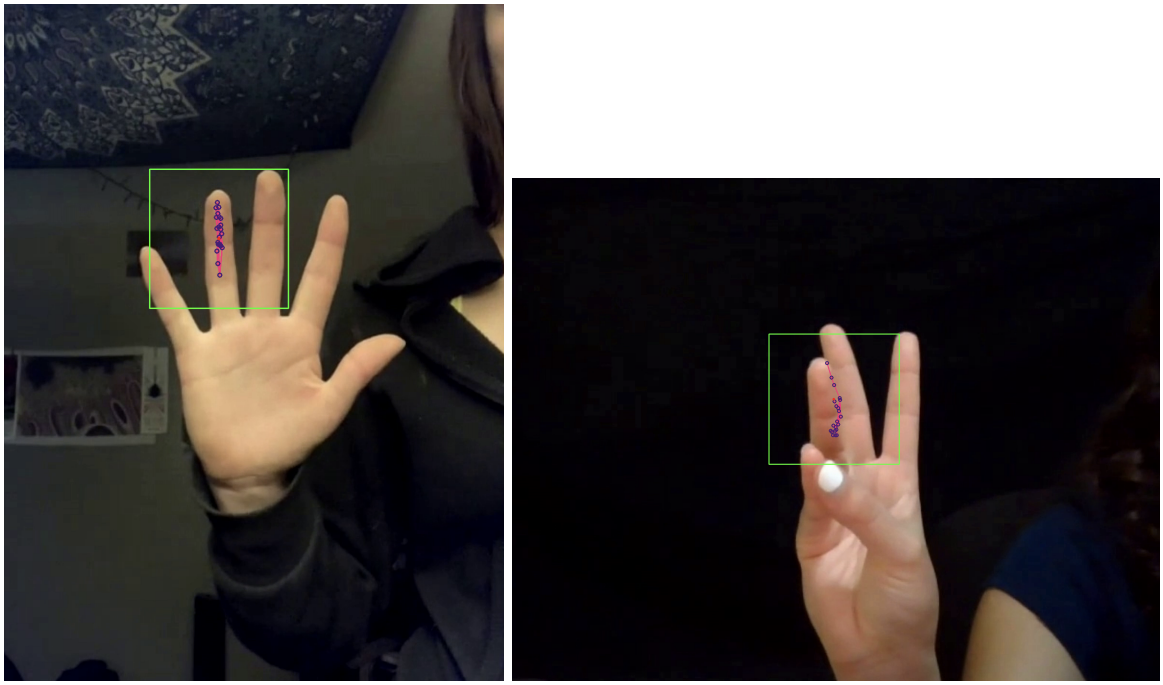


Figure: Frames that Mediapipe is unable to detect hand landmarks on, despite being clear, well lit, and in these cases with excellent contrast between the hand and background.

Overall, Mediapipe is meant to be a lightweight computer vision solution. It is meant for web and mobile applications such as this work, but it will make mistakes and will never be as fully accurate as “on board” solutions that need specific hardware and datasets, such as FaceID.

Motion detection when there is no skeletal information relies on a raster (aka pixel) based understanding of the frame. The frame becomes a matrix, which can then have background

subtraction performed on it such as when the convex hull of the hand centroids is found in the case of skeletal data as in section 4.6.

Background subtraction makes 2 key assumptions:

1. The background is more stagnant than the subject. This is because a still camera is assumed -- only active foreground objects move.
2. Grey world, which assumes that light is fairly constant within a threshold.

The first assumption is unfortunate because the camera will often move if people are recording on a phone or tablet. Additionally, some foreground objects move that are not hands, and are not subtracted as background due to this. For example, shoulders and faces.

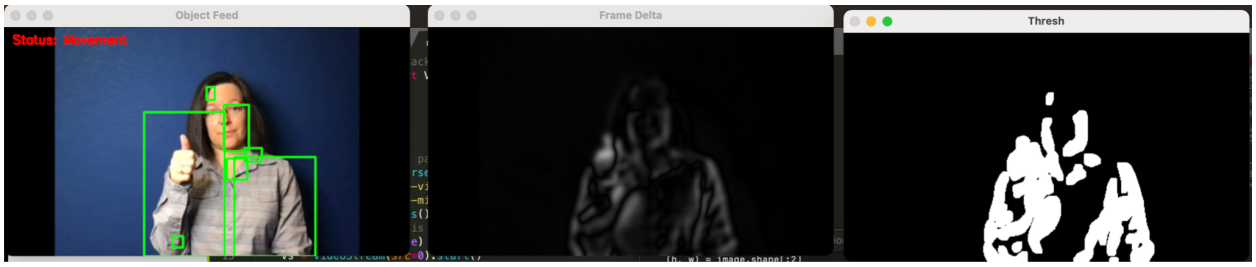


Figure: Thresholding of an example frame for motion detection experimentation

With regard to detecting and seeing hands for centering and utilizing a convex hull -- it is fortunate that it can be narrowed down to an ROI as covered in Section 4.

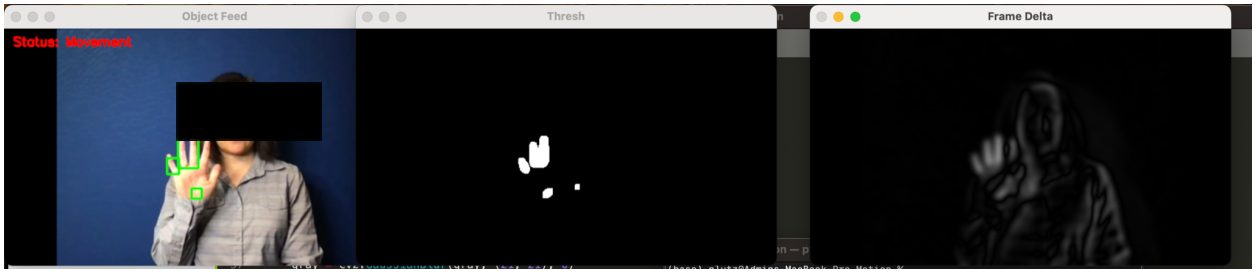


Figure: Improved thresholding for hands

Unfortunately though, for motion detection, there are performance issues. Unlike the geometric convex hull, the raster nature of the motion detection means that we need to keep track of more data points continuously to detect motion instead of keeping with particular local landmarks.

For example, one cannot simply just track the centroid of the ROI, as many signers will keep fairly still between signs, only moving fingers, which are as elaborated above, harder to track motion accurately with a moving camera and other noise conditions. And while the hand itself can be pre-run and isolated, there is still an issue regarding the amount of motion difference, especially with moving cameras and because convex hulls are approximations and it's very difficult to ensure accuracy around fingers in particular, which are the motion vectors needed to track for this slicing methodology.

This slows down the process and makes it harder to integrate overall, but it shows potential for future possible work.

6.6.2: Color and light balancing

Overall, due to the nature of these clips and that they are zoomed and scaled in, there is always going to be room to . The auto white balance and the normalization techniques outlined in this thesis are going to provide the best leveling for most neutral tones across a variety of skin colors and exposures. These balances make the videos more uniform across one another and across a variety of lighting conditions, cameras, and subjects.

Future development may go a step further and consider elements such as highlights, shadows, and exposure curves towards making more uniform edits to the footage samples.

6.6.3: Evolutionary improvements

Both the A Counting Spoken and A Counting Sign editions are generative artworks that evolve as more participants add their counts and transcribe other counts. These projects are and never were about translation or understanding the languages themselves, but rather about building dynamic systems for individuals to represent their language and create unified counts.

However, both systems have room for improvement. And there is a rich opportunity to create a pipeline that learns and improves from this to understand trends within these samples and identify possible issues with the segmentation or processing of this audio and visual information. This could improve not only these projects, but help to inform similar generative works.

6.6 4: Software improvements

It is important to note that this is not an engineering thesis or degree, but good software practices help enable future work in the intersection of art and technology. Overall, with any web application of this size, there are issues within the software and structure of it.

A few technical improvements for the codebase that enables this described thesis project may include but are not limited to:

- A formal test suite was never written for this software, and this could be a vast improvement towards edge cases and other future failures.
- There are relational database changes that could enable more flexibility and efficiency of the software.
- There are also points where the software could be more efficient or lighter weight.
- The video portrait itself could be made cleaner and simpler in implementation by writing the credits into the video instead of into an HTML page like the voice edition.
- The software could be made more robust against bots and other malicious actors.
- The code itself could be written more cleanly and modularly with more classes and a library like set up instead of a variety of batched processes
- Overall, there are likely more efficient and clean cut ways to do the centering, zooming, and panning of the videos for the final portrait. Both in the geometric interpretation of the process but also in the actual implementation.

6.6.5: Qualitative improvements

There is much to learn regarding this work and to improve it and future projects within these veins for them to better embody language and other justices.

A particular theme that emerged within doing this work is understanding the meta importance of language around the framing and execution of the artwork. To this point, current efforts regarding participation guides and materials for this project and others within the Poetic Justice research group are being developed. Understanding the importance of language around calls to participation and promotion and framing materials of artworks like these can allow them to reach more audiences and contribute to frameworks for future artists.

Ethically, this work could benefit from further examination and involvement of sign language users who would bring unique perspectives and find any potential issues or pitfalls regarding potential misuse but also any space for potential harm.

6.6.6: Potential future directions

This thesis implies a variety of frameworks and additional contributions in the line of this type of work. For example, there is much to be done in the context of the technical avenues explored from the color space and skin tone work to the segmentation of signs with motion detection and other computational methods.

Much of this work can and should be documented and open sourced in an accessible way for future developers to create more distributed video artworks like these on the web. The hope is that this thesis document will serve a portion of this future endeavour.

Artistically, there are many variations and experimentations that could occur in this artwork and more. What is an appropriate physical framing for a work like this? A screen? A projector? A physical booth? Is there a physical installation for recording that makes sense?

Learnings regarding the histories, scales, and relevant precedents in how they may relate to other language work in a more equitable and inclusive way are also important towards future development of conceptual and participatory artworks having to do with language and other identities.

7. Conclusion

There are several directions that this project could go in order to improve it on technical and qualitative levels. From an artistic perspective, there are a variety of ways this project could be displayed in physical and digital space. There are also a array of design possibilities that could be implemented to change the recording experience for participants, such as physical Such technical improvements, like segmenting the signs and improving performances, pose exciting new grounds for the technical depth of artistic projects such as these.

This thesis sought out to explore historical and artistic frameworks regarding participatory public artworks that engage with language justice, particularly signed languages and to document in great detail my contributions towards this space with *A Counting: Sign Language.* This proposes technical methodologies that can be used in a variety of related works, and also new qualitative frameworks for participation definitions in works like this and towards linguistic and language justice. And additionally, this shows some important tenets of Deaf culture and those who use signed languages.

This thesis is by no means exhaustive in technical, qualitative, or theoretical groundings. However, it hopes to serve as an example of rigorous historical and critical theory with extensive documentation of an art project that is enabled by technology. And to show the intricacies and intertwining contexts of how important language is to us all.

Sources

1. Lane, Harlan L, et al. "Origins of the American Deaf-World: Assimilating and Differentiating Societies and Their Relation to Genetic Patterning." Sign Language Studies, vol. 1 no. 1, 2000, p. 17-44. Project MUSE, doi:10.1353/sls.2000.0003.
2. Mitchell, Ross & Young, Travas & Bachleda, Bellamie & Karchmer, Michael. (2006). How Many People Use ASL in the United States? Why Estimates Need Updating. Sign Language Studies. 6. 10.1353/sls.2006.0019.
3. Bloom, Lary (2019-04-16). Sol LeWitt: A Life of Ideas. Wesleyan University Press. ISBN 9780819578709.
4. Grigely, J. (2016, November 1). Portfolio: Soundscaping. The online edition of Artforum International Magazine. https://www.artforum.com/print/201609/portfolio-soundscaping-64214
5. Native American Signed Languages, Jeffrey Davis, Subject: Linguistics, Signed Languages, Language Contact, Online Publication Date: May 2017, DOI: 10.1093/oxfordhb/9780199935345.013.42
6. Kim, Christine Sun. “I Performed at the Super Bowl. You Might Have Missed Me.” The New York Times, The New York Times, 3 Feb. 2020
7. Creative Capital. Kenny Dies and Allison O’Daniel. “Access in Content and Form”. Creative Capital. Virtual discussion. 2020.
8. Xu Bing. Website. xubing.com
9. Christine Sun Kim. Artist Website. christinesunkim.com
10. Chella Man. Website. chellaman.com/
11. http://karenmoss.art/touch/
12. “Assembly Line”. Website. https://arts.mit.edu/assembly-line/
13. “The Sheep Market”. Website. http://www.aaronkoblin.com/work/thesheepmarket/
14. “POSTCARD PROJECT”. Website. https://www.postcardproject-art.com
15. “SwarmSketch”. Website. http://swarmsketch.com
16. Hank Willis Thomas. Personal Website. https://www.hankwillisthomas.com
17. Arte Útil. Website. https://www.arte-util.org
18. Design Justice Network. Website. https://designjustice.org
19. Algorithmic Justice League. Website. https://www.ajl.org
20. Foster, David H. (2011). "Color Constancy". Vision Research. 51(7): 674–700. doi:10.1016/j.visres.2010.09.006. PMID 20849875. S2CID 1399339.
21. Almenberg, G. (2010). Notes on participatory art: Toward a manifesto differentiating it from open work, interactive art and relational art. AuthorHouse.
22. Linton, Simi (1998). Claiming Disability Knowledge and Identity. New York: New York University Press. p. 9.
23. Everett, Daniel (1988) On Metrical Constituent Structure in Piraha Phonology. Natural Language & Linguistic Theory 6: 207–246
24. Linton, Simi (1998). Claiming Disability Knowledge and Identity. New York: New York University Press. p. 9.
25. [Panel, 2020, Creative Capital, “ Access in Content and Form”, Kenny Fries and Alison O’Daniel]

26. Nora N Khan, “Second Readings”, 2020
27. Parkhurst, Stephen; Parkhurst, Dianne (1998), "Introduction to Sign Language survey", Ntes on Sociolinguistics, 3: 215–42.
28. arte-util.org
29. icp.org
30. mit.edu
31. ada.gov
32. nad.org
33. gdpr.eu
34. docs.opencv.org
35. gallaudet.edu
36. csd.org
37. Naik, N., J. Philipoom, R. Raskar, and C. Hidalgo. 2014. “Streetscore- Predicting the Perceived Safety of One Million Streetscapes.” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 779– 85.
38. Rethinking disability: the social model of disability and chronic disease, Sara Goering, Curr Rev Musculoskelet Med. 2015 Jun; 8(2): 134–138. Published online 2015 Apr 11. doi: 10.1007/s12178-015-9273-z.
39. Shingwauk Narratives: Sharing Residential School History. Jenna Lemay, Shingwauk Residential Schools Centre.
40. Bonvillian JD, Nelson KE, Rhyne JM. Sign language and autism. J Autism Dev Disord. 1981 Mar;11(1):125-37. doi: 10.1007/BF01531345. PMID: 6927693.
41. The Allan Sutherland Chronology – National Disability Arts Collection & Archive, https://the-ndaca.org/resources/the-allan-sutherland-chronology/.
42. Heroku.com
43. deafart.org
44. Redis task queues. redislabs.com.
45. docs.python.org
46. Lane, H.L., Pillard, R., & French, M. (2000). Origins of the American Deaf-World: Assimilating and Differentiating Societies and Their Relation to Genetic Patterning. Sign Language Studies 1(1), 17-44. doi:10.1353/sls.2000.0003.
47. Greenwald, B., & Van Cleve, J. (2015). "A DEAF VARIETY OF THE HUMAN RACE": HISTORICAL MEMORY, ALEXANDER GRAHAM BELL, AND EUGENICS. The Journal of the Gilded Age and Progressive Era, 14(1), 28-48. Retrieved August 10, 2021, from http://www.jstor.org/stable/43903056
48. When Color Constancy Goes Wrong: Correcting Improperly White-Balanced Images. Mahmoud Afifi, Brian Price, Scott Cohen, and Michael S. Brown, CVPR 2019
49. Deep White-Balance Editing. Mahmoud Afifi and Michael S. Brown, CVPR 2020.
50. W3.org
51. websitecarbon.com
52. nationaldeafcenter.org
53. webrtc-experiment.com
54. collab-project.github.io
55. vonage.com

56. opentok.github.io
57. linguisticsociety.org
58. tokbox.com
59. recordrtc.org
60. ncicap.org
61. cad.ca
62. developer.mozilla.org
63. tiktok.com
64. Memos, V.A., Psannis, K.E. Encryption algorithm for efficient transmission of HEVC media. J Real-Time Image Proc 12, 473–482 (2016). https://doi.org/10.1007/s11554-015-0509-3
65. Nyle DiMarco, IMDB.
66. Scarff, Larry A., "Quantification of the unsharp masking technique of image enhancement" (1981). Thesis. Rochester Institute of Technology.
67. Remarks on Color, Gary Hill, 1994, Artist Website
68. associationforpublicart.org
69. Jafa, A. Love is the Message and the Message is Death. High.org, Livestream. 2021
70. Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017, May 25). Photo-realistic single image super-resolution using a generative adversarial network. arXiv.org. https://arxiv.org/abs/1609.04802.
71. James Castle. Artist Website. jamescastlehouse.org
72. Buchmann, S. (2006). Art After Conceptual Art. MIT Press.
73. Bishop, C. (2012). Artificial Hells: Participatory art and the politics of spectatorship. Verso.
74. Bishop, C. (2006). Participation. London: Whitechapel.
75. K. Perlin, "Real time responsive animation with personality," in IEEE Transactions on Visualization and Computer Graphics, vol. 1, no. 1, pp. 5-15, March 1995, doi: 10.1109/2945.468392.
76. C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), 1998, pp. 839-846, doi: 10.1109/ICCV.1998.710815.
77. Spielmann, Y., Welle, A., & Jones, M. S. (2010). Video: The reflexive medium.
78. Hendren, S. (2020). What can a body do?: How we meet the built world.
79. Baker-Bell, A. (2020). Linguistic Justice: Black Language, Literacy, Identity, and Pedagogy. Milton: Taylor & Francis Group.
80. Millett-Gallant, A. (2015). Disabled body in contemporary art. Place of publication not identified: Palgrave Macmillan.
81. Siebers, T., & University of Michigan Press. (2010). Disability aesthetics. Ann Arbor: The University of Michigan Press.
82. Joseph Gigely, Bio FCA. .foundationforcontemporaryarts.org
83. Jones, C. A., Mather, D., & Uchill, R. (2016). Experience: Culture, cognition, and the common sense.
84. Kendi, I. X. (2019). How to be an Antiracist.

85. Language Justice, Antena. antenaantena.org/language-justice/. 2021.
86. Miller, Elizabeth, Little, Edward, & High, Steven. (2018). Going Public The Art of Participatory Practice.
87. Blosser, J. (2020). Relational History: Adam Smith's Types of Human History. Erasmus Journal for Philosophy and Economics, 12(2), 24–48. https://doi.org/10.23941/ejpe.v12i2.419
88. Spievack, N., & Gallagher, M. (2020, June 23). For Students of Color, Remote Learning Environments Pose Multiple Challenges. Urban Institute. https://www.urban.org/urban-wire/students-color-remote-learning-environments-pose-multiple-challenges
89. Semuels, A. (2021, June 14). The Online Hell of Amazon's Mechanical Turk. The Atlantic. https://www.theatlantic.com/business/archive/2018/01/amazon-mechanical-turk/551192/
90. Music . (n.d.). Brian Eno: 77 Million Paintings in Sarajevo. British Council Music. Retrieved August 11, 2021, from https://music.britishcouncil.org/projects/brian-eno-77-million-paintings-in-sarajevo
91. Waller, A. (2021, March 25). Black, Deaf and Extremely Online. The New York Times. https://www.nytimes.com/2021/01/23/us/black-american-sign-language-tiktok.html
92. Mesa, C. (2021, January 6). This Deaf Woman Did AN ASL Cover Of Cardi B's 'WAP' And When Is She Getting Hired To This For Shows? We Are Mitú. 100% American & Latino. https://wearemitu.com/fierce/this-video-of-a-woman-doing-asl-to-cardi-bs-wap-is-too-lit/
93. Hearing Aids: May 28 – Dec 4, 2021. (2021, August 5). McEvoy Foundation for the Arts. https://www.mcevoyarts.org/exhibition/hearing-aids-alison-odaniel
94. Teen who recorded Floyd's arrest, death wins Pulitzer nod. (2021, June 11). AP NEWS. https://apnews.com/article/pulitzer-prize-2021-citation-darnella-frazier-george-floyd
95. Cardwell, E. (2020, October 2). For Freedoms's Campaigns to Bring Creativity into Politics Are More Relevant than Ever. Artsy. https://www.artsy.net/article/artsy-editorial-freedomss-campaigns-bring-creativity-politics-relevant
96. Waterfield, S. (2019, April 15). ASL Day 2019: Everything You Need To Know About American Sign Language. Newsweek. https://www.newsweek.com/asl-day-2019-american-sign-language
97. Berke, J. (2020). How an Audist Attitude Negatively Affects Deaf People. Verywell Health. https://www.verywellhealth.com/deaf-culture-audism
98. Sutherland, A. disabilityarts.org. 2013.
99. Bragg, M. (2017, June 1). "The last remaining avant-garde movement." The Guardian. https://www.theguardian.com/society/2007/dec/11/disability.arts
100. Couch, C. (2020, April 2). Disability-Simulating VR Promotes Empathy. MIT Technology Review. https://www.technologyreview.com/2016/04/29/160444/disability-simulating-vr-promotes-empathy/
101. Tobias, J. (2020, August 18). The Tyranny of the Alphabet Book. NYARC. https://nyarc.org/blog/the-tyranny-of-the-alphabet-book

102. Kitowska-Łysiak, M. (2003, January). Roman Opałka. Culture.Pl. https://culture.pl/en/artist/roman-opalka
103. Davis, J. (2017, May 10). Native American Signed Languages. Oxford Handbooks Online. https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199935345.001.0001/oxfordhb-9780199935345-e-42#oxfordhb-9780199935345-e-42-bibItem-14
104. History.com Editors. (2019, November 30). Alexander Graham Bell. HISTORY. https://www.history.com/topics/inventions/alexander-graham-bell
105. Terrell, M. (2021, May 3). Alexander Graham Bell in Love. Travel. https://www.nationalgeographic.com/travel/article/alexander-graham-bell-in-love
106. Graber, J. (2018). Kiowa Sign Language. OUPBlog. http://languagehat.com/kiowa-sign-language/
107. Nović, S. (2021). The harmful ableist language you unknowingly use. BBC Worklife. https://www.bbc.com/worklife/article/20210330-the-harmful-ableist-language-you-unknowingly-use
108. Dipasquale, M. (2018, November 2). Don't Miss the Signs: Regional Accents in Sign Language. Publishing Solutions Group. https://publishingsolutionsgroup.com/blog/dont-miss-the-signs-regional-accents-in-sign-language
109. Lapiak, J. (2021). Accent in sign language. 1996–2021 Jolanta A. Lapiak. https://www.handspeak.com/study/index.php?id=3
110. Wilson, A. (2020, November 17). Life as a Coda: An Interview with Ku Mei Kern. GoReact ASL Blog. https://aslblog.goreact.com/life-as-a-coda/
111. Brown, K. (2020, December 15). 'I Want to Be Able to Maintain My Clear Voice': Artist Christine Sun Kim on Translating Her 2020 Into Trenchant New Drawings. Artnet News. https://news.artnet.com/art-world/christine-sun-kim-profile-1931118
112. Katz, S. (2021, February 26). The COVID Zoom Boom Is Reshaping Sign Language. Scientific American. https://www.scientificamerican.com/article/the-covid-zoom-boom-is-reshaping-sign-language1
113. Callahan, M. (2020, May 28). New challenges–and opportunities–for American Sign Language users during COVID-19. News @ Northeastern. https://news.northeastern.edu/2020/05/28/new-challenges-and-opportunities-for-american-sign-language-users-during-covid-19/
114. Player, D. (2021, July 2). Dear White Deaf People - David Player. Medium. https://dplayer84.medium.com/dear-white-deaf-people-490961afa3d4
115. Artist Christine Sun Kim on 'deaf rage,' the Super Bowl and the power. (2020, February 13). The World from PRX. https://www.pri.org/stories/2020-02-13/artist-christine-sun-kim-deaf-rage-super-bowl-and-power-sound
116. Gallaudet Cafeteria Mural and Chuck Baird. (2010, November 29). YouTube. https://www.youtube.com/watch?v=NFeHcrbn6iQ
117. Whisnant, S. (2019, March 7). Hidden history: How the education system overlooks harsh realities of Natives past and present. UNC Media Hub.

http://mediahub.unc.edu/hidden-history-education-system-overlooks-harsh-realities-natives-past-present/

118. Onion, R. (2021, February 26). Did Helen Keller Really “Do All That”? Slate Magazine. https://slate.com/human-interest/2021/02/helen-keller-tiktok-conspiracy.html
119. Lewis, S. (2019, May 31). The Racial Bias Built Into Photography. The New York Times. https://www.nytimes.com/2019/04/25/lens/sarah-lewis-racial-bias-photography.html
120. Steinhauer, J. (2014, December 11). Guggenheim Seeks Volunteer Readers for On Kawara Retrospective. Hyperallergic. https://hyperallergic.com/168016/guggenheim-seeks-volunteer-readers-for-on-kawara-retrospective/
121. Steinhauer, J. (2015, April 16). Reading Years for On Kawara. Hyperallergic. https://hyperallergic.com/199495/reading-years-for-on-kawara/
122. Frohne, U. (2013). Bruce Nauman – Ludwig Forum für internationale Kunst. LUDWIG FORUM FÜR INTERNATIONALE KUNST VIDEO ARCHIVE. http://www.videoarchiv-ludwigforum.de/artists-detail/bruce-nauman/
123. Powell, V. (2015, January). Image Kernels explained visually. Explained Visually. https://setosa.io/ev/image-kernels/
124. Bauman, H. et al. (2013, July). “BEYOND ABLEISM AND AUDISM: ACHIEVING HUMAN RIGHTS FOR DEAF AND HARD OF HEARING CITIZENS”. Canadian Hearing Society, Ministry of Education Report. https://www.chs.ca/sites/default/files/uploads/beyond_ableism_and_audism_2013july.pdf
125. Busch 2018. The language portrait in multilingualism research: Theoretical and methodological considerations. Working Papers in Urban Language and Literacies. https://heteroglossia.net/fileadmin/user_upload/publication/busch18._The_language_portrait_copy.pdf
126. Beaulieu, D. 2009. “26 Alphabets (for Sol LeWitt)”. https://derekbeaulieu.files.wordpress.com/2016/08/26_alphabets__for_sol_lewitt_.pdf
127. McGregor, T. 2003. “Keresan Pueblo Indian Sign Language”. Nurturing Native Languages (pp. 141-148). Flagstaff, AZ: Northern Arizona University.
128. The Question of Sign-Language and the Utility of Signs in the Instruction of the Deaf: Two Papers by Alexander Graham Bell (1898) The Journal of Deaf Studies and Deaf Education, Volume 10, Issue 2, Spring 2005, Pages 111–121, https://doi.org/10.1093/deafed/eni012
129. mtsac.edu. “5 Parameters of ASL”.
130. “Stoke Notation Introduction.” Julie A. Hochgesang for Linguistics 101 at Gallaudet University created fall 2007, updated August 2015.