Translations

Designing Restorative Listening Experiences in the Age of Social Fragmentation

by

Naana Obeng-Marnu

B.A., Brown University (2017)

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 in Media Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2024

Author_____

Naana Obeng-Marnu
Program in Media Arts and Sciences
May 17, 2024

Certified by_____

Deb K. Roy
Professor of Media Arts and Sciences
MIT Center for Constructive Communication
MIT

Accepted by_____

Joseph A. Paradiso
Associate Professor of Media Arts and Sciences
Program in Media Arts and Sciences
MIT

Translations

Designing Restorative Listening Experiences in the Age of Social Fragmentation

by

Naana Obeng-Marnu

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on May 17, 2024, in partial fulfillment of the
requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

This thesis builds on a body of sociotechnical research at the MIT Center for Constructive Communication that draws upon "ancient wisdoms" of dialogue and listening and harnesses the power of technology to inform the design of dialogue spaces that promote deep, meaningful, and authentic conversations. Our approach hinges on the belief that society functions best when we hear and understand each other, an outcome that our work strives to advance by exposing people to the personal stories of others in ways that connect rather than divide. I take inspiration from anthropological practices and recent Data Humanism and Activism epistemologies to derive a set of design considerations for restorative interfaces. These principles inform the development of *Translations*, an interactive experience that invites audiences to more deeply engage with a curated collection of stories surfaced during small group facilitated conversations. The design of this visual and auditory experience allows audiences to explore stories they may otherwise not hear through websites that center thematic summaries and high level insight visualizations. The selected stories are curated using AI emotion analysis and sensemaking which are leveraged to draw the user's attention to moments of interest across conversations, such as moments of affirmation. The efficacy of this curation method to engender empathy and emotional disruption, precursors to restorative listening, is evaluated and the results from user tests for and interviews about the overarching interface are discussed. Ultimately, this thesis presents both a framework for automatic curation of audio narratives as well as an interactive interface to present these selected stories, both of which have wide-ranging applications in the media and civic space.

Thesis Supervisor: Deb K. Roy
Title: Professor of Media Arts and Sciences, MIT Center for Constructive Communication

Translations

Designing Restorative Listening Experiences in the Age of Social Fragmentation

by

Naana Obeng-Marnu

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

Thesis Reader_____

Deb K. Roy
Professor of Media Arts and Sciences
MIT Center for Constructive Communication
MIT

Thesis Reader_____

Erica C. James
Professor of Medical Anthropology and Urban Studies
Department of Urban Studies and Planning
MIT

Thesis Reader_____

Zach Lieberman
Adjunct Associate Professor of Media Arts and Sciences
Program in Media Arts and Sciences
MIT

Thesis Reader_____

Dimitra Dimitrakopoulou
Head of Translational Research
MIT Center for Constructive Communication
MIT

Acknowledgements

I would like to acknowledge the following people, communities, and institutions without whom this thesis would not have been possible. I truly stand on the shoulders of giants:

First, I would like to thank my advisor Deb Roy for his guidance and mentorship and abundance of patience throughout this process. You provided a unique perspective through which I could view my work, allowing me to consider possibilities previously beyond my imagination. I would also like to thank Dimitra Dimtrakopoulou, Head of Translational Research at the MIT Center for Constructive Communication (CCC) and my secondary advisor within the research group for your continuous support, sacrifice, and guidance which deepened my research methods.

I would also like to thank my other thesis committee members and readers Zach Lieberman and Erica Caple James. Both of your courses which I took during my first year in the program inspired and informed much of the work in this thesis, and, were it not for your many literature and artistic references, I would be lost. Erica, I have truly enjoyed every meeting and correspondance. Your expertise has been a beacon of light. And, Zach, your Drawing++ course truly changed my life and started me on this data art journey. You have made an entire field of creative coding accessible. I am truly indebted to my dedicated committee who sacrificed to sprint towards the finish line with me. Your feedback and attentiveness were invaluable.

I would also like to thank the professors and instructors of other courses I took at MIT whose tutelage has allowed me to grow my artistic, design, and research parctices: Pattie Maes, Caitlin Morris, and Abhinandan Jain (Extended Cognition), Danielle Wood (Space Technology for the Development Leader), Mitch Resnick (Learning Creative Learning), and Gearoid Dolan (Introduction to Interactive, Participatory, Generative Art Making).

I would also be remiss if I did not thank the many students and researchers who consulted on and assisted me with various components of the thesis. To Char Stiles, thank you for your diligent explanations of and tutorials on shaders. To Perry Naseck, thank you for consulting on the back end user tracking. Thank you Daniel Kessler for your guidance on evaluation models and various survey scales. Thank you also to Hope Schroeder and Wonjune Kang for your feedback and recommendations on my study design. Thank you to Cassandra Overney for your general thesis advice. Thank you to Tobin South and Marcus Edmonds for your assistance with ChatGPT and for challenging my perspectives on AI despite my reluctance. Thank you Alex Berke for your general advice and mentorship on various aspects of my thesis and research practice. Thank you also to Brandon Roy, Titi Phan, Jad Kabbara, and Dennis Jen for your assistance with the NLP and AI Sensemaking components of this thesis. Thank you to Niko McGlashan, a wizad and knight of code, who chivalrously helped me troubleshoot various problems that arose. And lastly, but certainly not least, thank you to Vera van de Seyp, goddess of p5.js, for elevating my p5 skills and advising on the design.

Thank you as well to the experts that advised me during the course of this work: Graham Jones, Catherine D'Ignazio, Chris Bannon, Peter Litman, Phil Chang, and Elise Hart Kipness. Thank you to the staff at Cortico without whom this thesis would not be possible. Thank you Georine Pierre, Nicole Bakker, Jessie Mindel, and the rest of my Media Arts and Sciences Master's cohort. Thank you aslo to the Media Lab's Student Committee (STUDCOM), my co-editors of SEEDVAULT, and 99Crew, especially Eyal Perry, my STUDCOM work husband. It has been an honor to explore the bounds of technology, creativity, and play with you all.

I would also like to thank the friends and family who have supported me throughout this process. Thank you Mom, Dad, Kwadwo, and Abby for your continuous encouragement. Thank you also to the Kipnesses: Elise, Rob, Justin, and Ryan for your moral support and the beautiful haven you have established in your home. Thank you Al Weiner, Marcus, Tobin, and those who have resided in Calvin house these last few years. Your home is my home. Thank you Shayne Longpre and Rob Mahari for the opportunities you provided me to hone my visualizations skills through the Data Provenance Initiative. Thank you to Kash Singhal and Cristian Mendez for being incredible beta testers and sounding boards. Thank you to my roommates Julian O'Donnell, Alexa Thompson, and Rosie Keller. And thank you to friends Kush Tiwary, Suyash Fulay, Paris Meyers, Phil Chang and his wife Jessica, Francesca Agege, Marina McNamara, Patricia Burden, Jade Lu, McCoy and Veronica Bean, and, of course, I could never forget Blake Ridley.

Thank you to Cassandra Lee, Maggie Hughes, and Belén Saldías Fuentes, my partners in crime at CCC. You have been incredible collaborators, mentors, and friends. I have truly appreciated all that you have done for me and all that we have accomplished together these last two years. There is so much I could say about you all, my Ars Angels, but I will just share that I know partnerships like ours are rare, and I have never taken you for granted.

Last but not least, I'd like to acknowledge and thank everyone at CCC, including many I have already mentioned and current and former members such as Russell Stevens, Heather Pierce, Leila Carter, Matthew McKenna, Marina Rakhilin, Maya Detwiller, Artemesia Luk, Hang Jiang, William Brannon, Shrestha Mohanty, Michael Wong, Elinor Poole-Dayan, Matt Nay, Sam and Annelli Woolf, Leonard Francis Vibbi, Bridgit Mendler, Eugene Yi, Andrew Heyward, Debby Goldberg, Ellen Hoffman, Lorena Tovar, Lorrie LeJeune, Rukun Zhang, Sinclair Target, Wes Chow, Doug Beeferman, Nabeel Gillani, and Amy Johnson. It has truly been an honor working with you all. I am a better person and researcher for knowing you.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the wake of rising hate speech on what was once known as Twitter following Musk's acquisition of the platform, there have been reports of a significant portion of its black users fleeing or maintaining their accounts but going silent [1]. This vacuum has sent many on a quest for new communication spaces and has resulted in the development of competitors for the legacy social media site [1]. This recent digital diaspora mirrors historical disruptions to both physical and media landscapes, following which, affected communities developed new vocabularies, practices, designs, and homes that became foundations for commiseration, action, and restoration.

Culture, politics, and activism thrive when people are given the opportunity to commiserate, laugh, and collectively make sense. Indeed, the appeal of Black Twitter, in particular, was the community unlocked through linguistic codes (hashtags), visible to all who knew where and when to look but only accessible to those adept at these codes. The demise of such a black landscape would result in a loss of collective memory, belonging, allyship, and sensemaking. But, as Walter Hood states in his introduction to Black Landscapes Matter, "Erasure is a call to arms to remember" [2]. While there are those who seek to archive Black Twitter before it is lost [3] or attempt to reclaim it on similar social media platforms [4, 5], we can instead imagine, as we have in the past when faced with erasure, new landscapes and cultures, strengthened by the affordances of recent technologies, where we can better

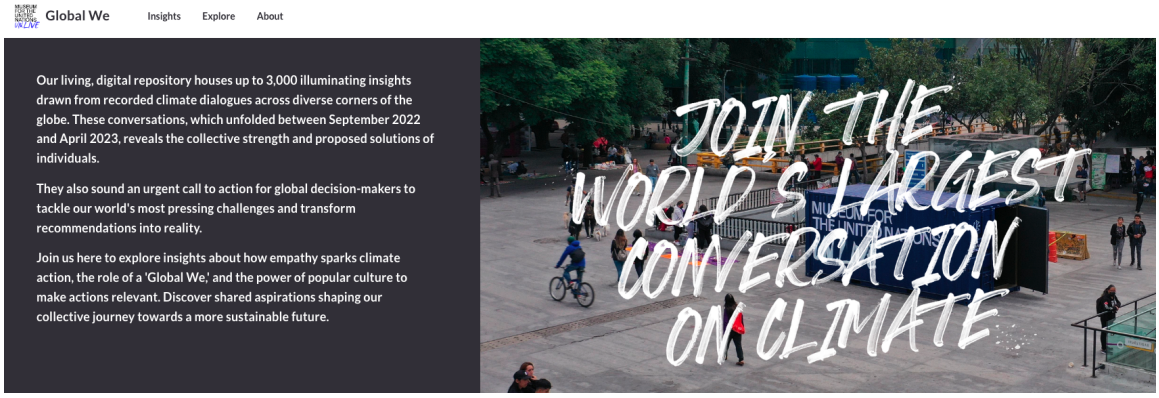

Figure 1-1: Screenshot of the landing page of the Global We conversation portal. The Global We is a conversation collection organized by the Museum of the UN.

explore our collective memory and make sense of our shared and unique experiences.

Civic Tech researchers are in the process of designing such a reimagined digital public square. One such group, the MIT Center for Constructive Communication (CCC), of which I am a part, has put forth an alternative restorative practice: a social dialogue framework, in contrast to the social media networks that prioritize broadcasting over listening [6]. These methodologies and tools, supported by recent advancements in natural language processing and other technologies, rely on facilitated, small group dialogue and collaborative sense-making processes that allow for more productive conversations to occur across difference. In this framework, communities are also given the opportunity to make sense of and explore insights from these conversations, engaging with the power of collective experience. It is on this part of the process that I will focus.

In the past, insights formed through this sensemaking process have been shared through online interactive websites known as conversation portals and used as prompts for further discussion such as conversations with mayoral candidates in Boston [7]. But, questions remain about how we can further engage the public in listening: how we can develop an experience that allows them to connect with their neighbors' stories even if they themselves did not participate in one of the small group dialogues? How can we allow them to see themselves reflected in our memory pool–or flag when they are not–to empathize with their ideological rival even if they do not agree, and to be heartened to continue to participate in

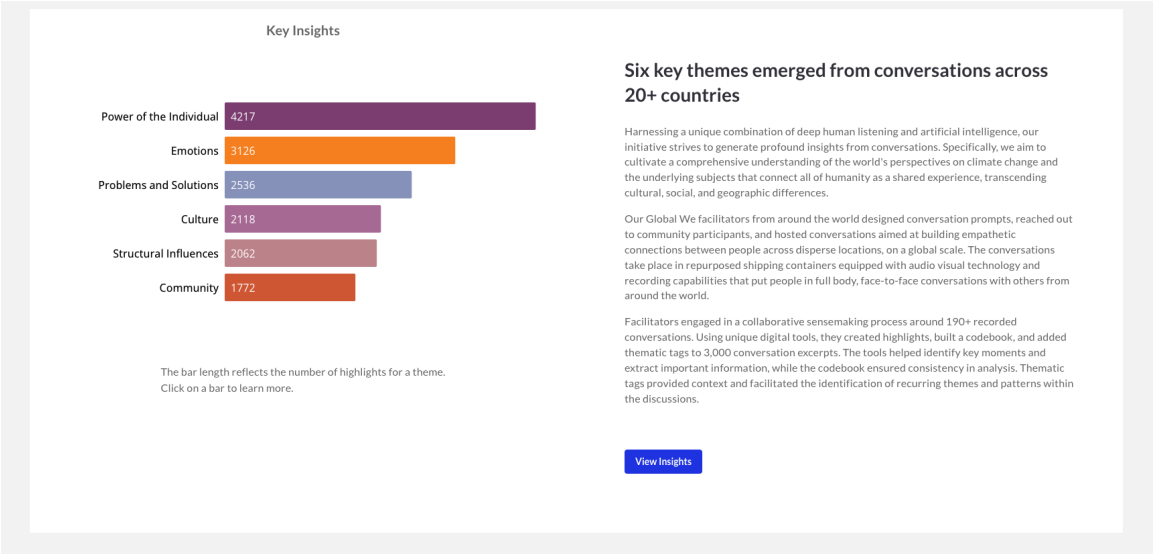

Figure 1-2: Screenshot of key insights from the Global We conversation collection as presented in its conversation portal.

civic discourse? And for those tasked with making decisions on behalf of these communities, how can we encourage them to spend more time and empathize with the voices that have been collected and stored in our virtual conversation library and identify issues of importance?

In this thesis, I look beyond "ancient social technologies"–a label presented by Hughes [8] to represent long-standing techniques of eliciting fruitful conversations such as circle practice–to visualize the practices we employ that can serve as indicators of moments of significance to be leveraged to engage others. I speak of the "'we' experience" [9] (how we respond to one another and how ideas and values emerge from our collective experience); the "everyday mundane" and "the cultural turn" embraced by Hood [2] (how we create space for and represent the uniqueness of ourselves and relate to the humanity of others); and sacred shouts, collective cries, "a sacred attention" (how we commiserate, empathize, and reflect) [10]. These can be translated to practices in the digital space such as grounding stories in context and limiting the field of view to direct attention, for example. We can learn from these community practices to inform the design of new spaces–spaces that draw attention to not only the dominant voices but also shared experience and collective context.

Through Fora, an online platform that hosts community conversations maintained by the

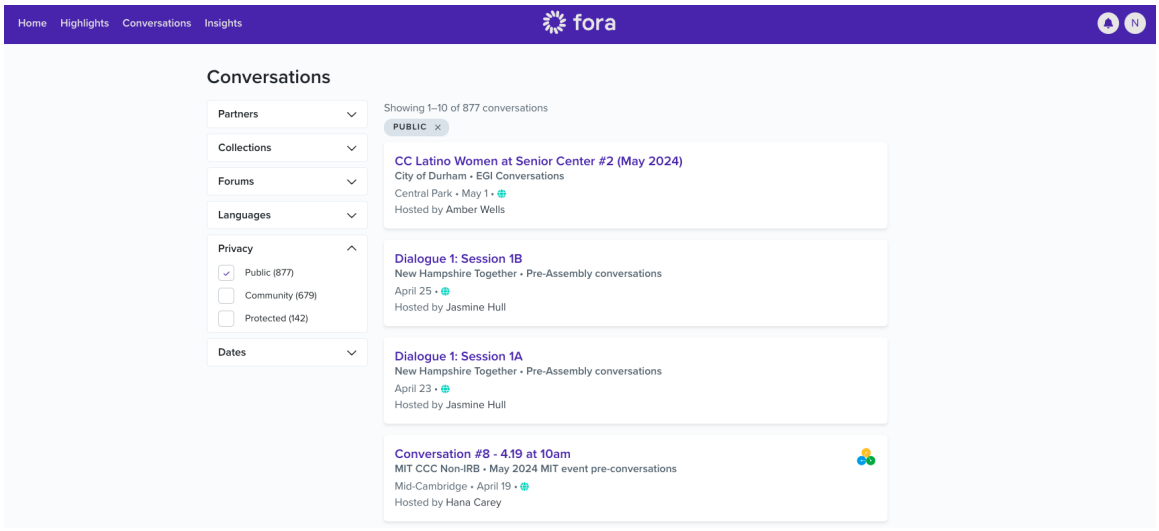

Figure 1-3: Screenshot of the Conversation search within Fora Desktop. Users can filter and search for conversations that match parameters such as desired collection and privacy level.

non-profit Cortico with which CCC cooperates, CCC has access to an extensive catalog of facilitated community dialogues. These conversations were held for a myriad of purposes from soliciting input regarding candidates for civil service to workplace mediations [11]. This digital conversation library provides an opportunity to analyze and visualize community conversations. It is the subset of local conversations within the library held for the purpose of soliciting and elevating community stories to inform civic processes and policies on which I will focus for this thesis. The participants of these conversations shared their testimonies with the hope that their values, affirmations, and concerns would be heard by local civic leaders, and it is my intention to design experiences that allow the target audience as well as community members who may not have participated in the conversations to engage with these stories.

Thus, the guiding question in this work is, How might automated, systematic curation, interaction design, and principles from related disciplines such as anthropology inform and encourage deep listening? This overarching question led me to the following research questions:

- **RQ 1**: How might we surface and visualize moments of significance, consensus, and

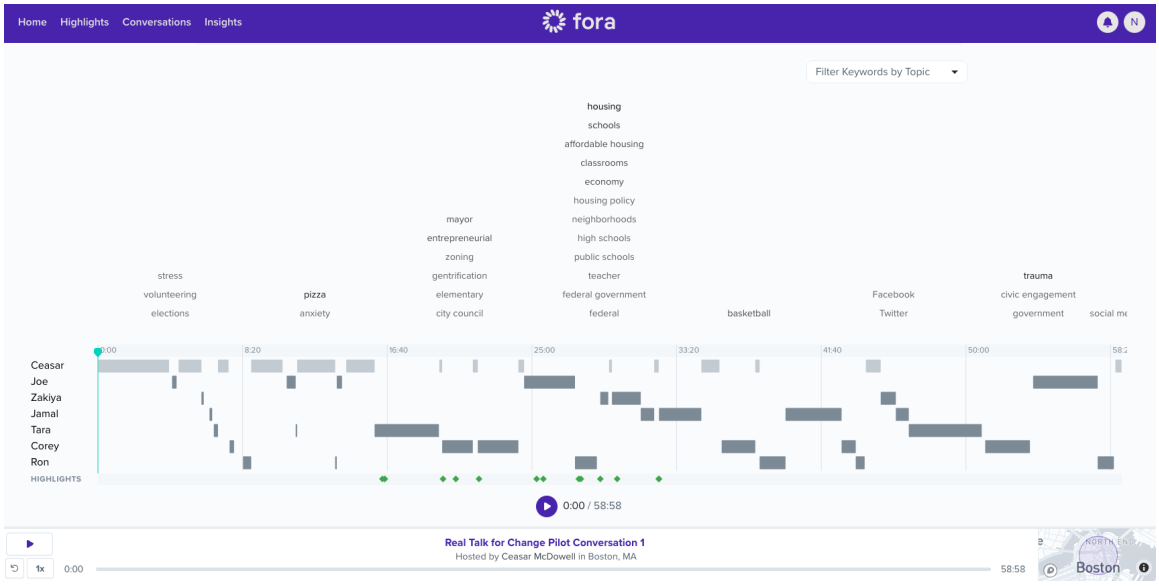

Figure 1-4: Screenshot of the interactive transcript "ribbon" within Fora Desktop. Speaker turns are visualized across time and by speaker. Users can click on a speaker turn to play the associated audio or select themes to see related turns.

tension in a conversation corpus? How do we identify and assign weight to these moments, given the conversation audio as input? What cues can be derived from the dynamics and utterances of the conversations' participants?

- **RQ 2**: How would an interactive interface populated with such audio-visual curation mediate a user's connection with the corpus?

I hypothesize that (**H1**) presenting a user with stories curated through a process that leverages key social signals (emotions and utterances we wish the end listeners to experience) will indeed engender connection with the stories and those who shared them. Additionally, (**H2**) my proposed interface design will guide the listener to slow down and more actively engage with the corpus than the standard thematic visualizations and summaries often presented in the conversation portals or the interactive transcripts and timelines found on the Fora website, which often lead to passive listening. And, it is this activity that will further facilitate connection with the speakers and their stories.

The key objectives and contributions of this thesis include:

- A systematic method of curation, informed by the reactionary utterances of the conversation participants.
- An interactive audio-visual interface that allows listeners to explore this pool of voices, navigating the stories that reflect the interests of a conversation collection's participants.

I title this work *Translations* because I aim to translate the collective experience and sentiments of communities to third-party audiences. A third-party public may not have the tools to understand and empathize with the experiences shared by a given community but my hope is that through this audio-visual framework they may begin to hear and desire to understand. They will see at a glance not only what was talked about but how the community felt. This work is not meant to replace traditional sensemaking–the manual work of listening to conversations in full and tagging key excerpts with relevant codes–or summarization but explores how recent technologies and epistemologies may be used to draw attention to moments of interest that may encourage further exploration of a full corpus. In fact, I will demonstrate that the explorations initiated by my tool incite curiosity and would lead to deeper engagement, exploration, and annotation. Harding posits that conviction "comes only through speech" [12]. It is through a careful selection of words that conviction and connection are espoused. Active listening, participation, and engagement can bolster this effect. Digital platforms are good at making a version of collective experience visible but active listening is difficult in digital space where we are all susceptible to mindless scrolling. Through *Translations* and my related work, I envision alternate modes of engagement.

In the coming chapters I will ground my work in related theory and practice before diving into an explanation of the framework, the development of the visual language and interface. Before finally, evaluating its efficacy.

Chapter 2

Background

The MIT Center for Constructive Communication (CCC) seeks to design spaces that promote civil, empathetic, and productive conversation, surfacing voices from underheard communities through facilitated dialogues. This mission was inspired by prior work from the Center's previous incarnation as the Laboratory of Social Machines (LSM) research group at the MIT Media Lab [13]. Among other research areas, the LSM group studied social dynamics on traditional and social media platforms and uncovered concerning polarization trends and echo-chambers of discourse [14,15] that existed in these spaces. Siloed communities in these virtual and over the air environments precipitated fracturing in physical social spaces [16–18]. To put it simply, after nearly a decade of communities talking to like-minded individuals, their ability to empathize and communicate with those who differed from them eroded [17,19]. Around this time, LSM, in cooperation with the non-profit Cortico, launched the Local Voices Network, a platform that hosts community conversations organized and facilitated by local community-based organizations [20]. In-person conversations were held around a recording device called the Hearth [21], also developed by Cortico, which uploads the recordings to their LVN platform for semi-automated transcription and diarization as well as human-led qualitative coding through a process known as sensemaking [22–24].

In the years to come, the work of LVN became a framework for listening and advocacy. MIT CCC and Cortico have since scaled the project to several partners across the globe. The

Figure 2-1: Photo of version 2 of the Hearth by Cortico. The Hearth is the recording device used to collect the in-person accounts of those who participate in a Fora small group, facilitated dialogue.

listening and recording mechanisms may now vary by partner or collaborator (hereafter referred to as the conversation "host"), location, and objective, with conversations occasionally hosted over Zoom or even in shipping containers [25,26]. And, recently, LVN has relaunched as Fora, a suite of tools that makes the principles championed by CCC and Cortico widely available and accessible [27].

However, their sensemaking process [22] remains largely intact: following transcription, a group of human sensemakers made up of nominated members from the host organization with support as needed from Cortico's partnerships team and/or CCC researchers develop a codebook of themes of interest then highlight and tag key excerpts from the transcripts. Thematic summaries may also be written by the sensemaking team. For some partners and initiatives, an online website ("portal") populated with the summaries, chosen highlights, and accompanying visualizations of thematic trends may be built to allow the community within which the conversations were held and perhaps even the greater public to explore the insights gleaned from these conversations [7,25,28]. Conversation collections vary in size–the Real Talk For Change Boston collection on which much of my research has been centered

contains 72 conversations while the largest collection The Global We for Climate Action consists of 331–and require a significant manual lift from all those involved [11]. Thus, CCC continues to research methods to leverage emerging technologies to improve and enrich these processes.

Prior work in this group, such as Hughes' Keeper, has employed visual interfaces to scaffold (1) the facilitation of such dialogues as they unfold [29]; (2) the qualitative coding and thematic analysis of these conversations [30]; and (3) the sharing of these insights back to the relevant communities [31]. Meanwhile, other efforts in the group include leveraging recent advancements in Natural Language Processing to incorporate AI in the loop for processes such as sensemaking [30] and the development of new metrics and visual languages to reveal and interpret conversation quality [32, 33].

Still, much can be done to interrogate how stories are shared. The goal of the group is ambitious: to bolster democracy through constructive communication. The process of facilitated dialogues and sensemaking may be effective for those who are able to participate, and exposure to personal experiences rather than opinion engenders empathy [34,35]. But how do we grant these experiences to those who cannot be in the room (or Zoom)? One way, on which the group is currently working, is to scale this service to a social app, bringing the tools and strategies of the social dialogue network to everyone's pocket [27]. Fora mobile would allow a wider audience to experience the social dialogue network first hand and engage in these conversations wherever they are. This social dialogue networking app would set itself apart from traditional social media apps by (1) allowing users to participate only in small group dialogue (2) giving these users the opportunity to do their own sensemaking of the conversations they participated in, and (3) allowing only the highlights approved by those whose voices are included in them be shared through the various thresholds of privacy, from community-level to fully public. But, for those who may still be hesitant to exchange stories in such a way, another method of reaching them is to make listening to existing libraries more accessible. Allowing opportunities of exploration of the existing corpus where possible models the behavior we wish global citizens to embody. And, by seeing the positive effects, they may be encouraged to participate down the line.

2.0.1 Ethical Loneliness and Restorative Listening

In her 2015 book *Ethical Loneliness*, Jill Stauffer describes the phenomenon after which her book is titled as "the experience of being abandoned by humanity compounded by the experience of not being heard" [36]. Throughout the book, Stauffer illustrates the plight of those abandoned by society–the unheard who are twice victimized, first by their oppressors and then again by institutions and communities who refuse to or are incapable of hearing them. There are many reasons Stauffer cites that may prevent one from hearing, the irony being that in several of these cases, it is not always that the institutions or individuals tasked with listening do not wish to listen. In fact, many believe that they are. In one example she presents, interviewers of the Fortunoff Video Archive for Holocaust Testimonies at Yale University attempt to collect the account of a Holocaust survivor Hanna F. but unwittingly silence the woman following mutual misunderstandings. One interviewer calls her "plucky," referring to her audacity to defy Nazi officials three times to ensure she made it on a bus leaving Auschwitz. It appears Hanna at first misheard the interviewer, thinking they said "lucky," so she refuted their term, stating it was stupidity not luck. The interviewer amends Hanna's correction to add that it also took "guts." At thus point, Hanna presumably gleans the interviewer's intention but still disagrees with their assertion. She rebuts their understanding of her story, saying "there were no guts, there were just sheer stupidity." The interviewer stands their ground. Hanna tries again to explain her point, but another interviewer stands up and ends the interview. The interviewers in this case are unable to hear her because she is describing an experience that is not aligned with their expectations. In this and other examples, Stauffer shows us the intention of listening is not enough to address ethical loneliness even when one is sitting across from someone. The story the listener tells themselves must be disrupted before they can hear and before the speaker can be heard.

I draw inspiration and caution from Stauffer's work because we live in times fraught with ethical loneliness, and it is this phenomenon that I wish my work may alleviate. To further illustrate this, I will share another example from my own experience. In 2021, amidst the pandemic and following the tragic death of George Floyd, several companies including my

own employer at the time Meta, held conversations in order to listen and glean insight into the black experience. I attended several meetings where black coworkers were meant to share their pain. But, oftentimes peers and leadership only wished to hear experiences that reinforced their ideas of themselves–that they were not part of the problem. Their tears outnumbered those of the people who had experienced the harm. This was not truly listening. This was not empathy. The result of these conversations was an additional day of training on Juneteenth which none of their black employees had asked for or needed (at least not in the manner it was implemented). This was not restorative justice.

Here, restorative justice refers to the practice championed in criminal justice and education spaces of facilitating accountability and rehabilitation through communication [36, 37]. It is restorative justice that ethical loneliness calls for. Thus, to remedy the ails of modern society we must invent restorative listening spaces, developing tools that allow us to hear and hold each other accountable so that we may begin to rebuild. How we design such listening spaces is the primary quest of this thesis. As Stauffer asserts, listening requires disruption, a challenge to the stories we tell ourselves. This disruption is often uncomfortable and can lead to emotional arousal. Listeners may be inclined to use emotions even as a defense. However, restorative listening also requires empathy. Through *Translations*, I provide an opportunity for empathy.

Modes of Listening

There is a longstanding tradition of qualitative research, interviewing, and facilitated dialogue as a method of listening [38, 39]. In service of such listening, qualitative researchers use tools such as NVivo to understand and interpret the stories of their subjects [40]. And, as previously mentioned, Cortico's Fora offers a qualitative coding framework that allows hosts and sensemakers to review and annotate stories [11]. These are functional tools that allow for the elevation of community experiences but are limited to the professional and academic space. Tools for the public to engage with such corpuses and make sense of their communities are limited.

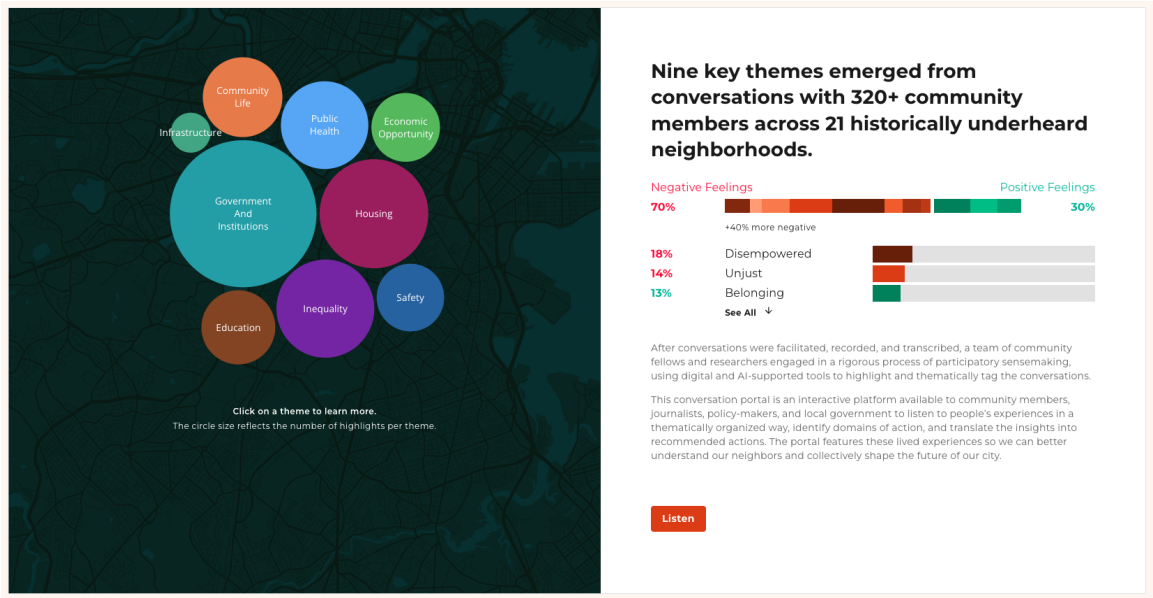

Figure 2-2: Screenshot taken from the landing page of the Real Talk for Change Boston portal. To the left is a circle packing chart of the themes that appeared across the collection. To the right is a summary of the emotion breakdown of the collection with a brief description of the sensemaking process.

Within the Fora framework, following the sensemaking process, insights are collected and presented in online portals. These offer the communities in question and the greater public the opportunity to review the themes that are in the conversation by reviewing the visualizations and summaries presented. Users can listen to selected excerpts from the conversations as well. However, when navigating these portals, though I can see the themes discussed and the volumes of stories within each theme, I am often left curious about the participants' experiences while sharing the story. Are these the stories they wish us to hear? In addition, due to the sheer volume of information presented, it is difficult to resist the urge to passively browse–click on a handful of stories and exit the page. Though the summaries and insights are grounded in the original conversations–excerpts are cited and the user is invited to play the contextual audio to verify the source of the claim–many may be inclined to gloss over these. In this scenario, users learn about the conversations at a high level but they may not engage with the voices. This experience is not equivalent to the participants' listening experience. In this case, the third party user does not feel what it was like to be in the room, and thus, do not fully benefit from the principles of the social dialogue framework.

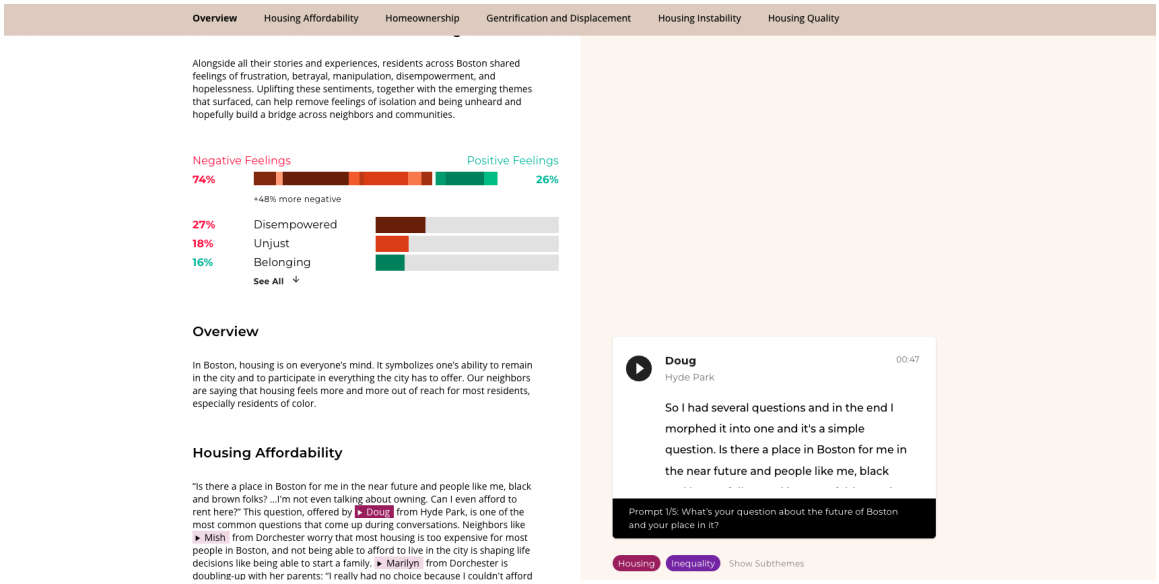

Figure 2-3: Screenshot of the insights and summary page for the theme "Housing" taken from the Real Talk for Change Boston Conversation portal. Playable audio excerpts are included within the summary akin to citations.

So, how do we encourage listening and engagement? And how do we engender empathy and connection to the stories in this corpus?

First, we must consider context. I have made the claim that the third party experience lacks the necessary context for connection. The listener requires an understanding of participant experience, but this is difficult to simulate outside of the conversation. How does one capture collective, narrative experience? How does one translate context and emotion? This problem is not new: anthropologists and other such purveyors of oral histories have grappled with this issue as well. For this reason, I turn to these practitioners for inspiration. In Antoinette Errante's 2000 paper "But Sometimes You're Not Part of the Story: Oral Histories and Ways of Remembering and Telling," she highlights the importance of representing the collective voice in anthropological accounts of oral histories [9]. She presents the metaphor of a "memory pool," emphasizing the necessity of translating the full context of a conversation: "The 'voice' narrated in these stories was a 'collective' one; it was a 'we' experience and perhaps for this reason, it required a 'we' context in order for people to narrate it" [9]. From this and similar studies, I surmise the importance of being present in the room and

observing the subtle social exchanges that exist as the conversation unfolds. So, perhaps, if we capture and translate these social exchanges, the listener will be encouraged to linger and empathize.

Next, we consider the disruption Stauffer describes. If we wish listeners to take action or to at least be moved, what is required? For this I turn to Susan Harding's 1987 journal article, "Convicted by the Holy Spirit: The Rhetoric of Fundamental Baptist Conversion." In this essay, Harding recounts her experience with a Baptist preacher on her quest to understand how such figures are able to incite "generative belief": "belief that indisputably transfigures you and your reality" [12]. After her encounters with the preacher, Harding discovers that she herself has been swayed somewhat by his words. Harding explains that what is necessary for conviction is (1) active and uncritical listening and (2) the vulnerability in the exchange of intimate storytelling [12]. What swayed her was the pastor's willingness to share and hers to participate in the exchange.

We would expect others to go through similar transformations after engaging in such an experience. However, facilitating such experiences in the digital space encounter two hurdles: (1) how does one actively listen in virtual space and (2) with a pre-recorded corpus, how do we translate the vulnerability of intimate sharing. I assert thoughtful motion and interaction design can address the first hurdle and careful curation and visualization address the second. These techniques fall under the umbrella of artistic representation.

2.0.2 Drawing Dialogue

My work is situated against a backdrop of conversation drawings and visualizations, a legacy of artists and researchers capturing various dimensions of human social interaction in ways that reveal and delight. This includes more technical visualizations such as Chandrasegaran et. al.'s TalkTraces which visualize meeting records in real time [41] as well as Judith Donath's efforts to illuminate the characteristics and potential harms of digital social interaction and publicly displaying affiliation [42]. Through the blossoming flowers of People Garden, Donath and her collaborators capture dominance while they suggest social fragmentation

and isolation in Loom [43].

Yet, I would often say that this thesis is not a visualization project but a curation, design, and storytelling project. The visual language and metaphors are a core piece of how the components may be interpreted, but,–to use the tenants of visualization presented by Donath et. al.–beyond data choice, generation, and mapping, I was most concerned with impact [44]. I am drawn to Auguston Nagy's work which prioritizes abstract expression of the subjects he visualizes [45]. Discernable patterns may not be gleaned at first glance, but his audiences are captivated. Then, upon closer inspection, one can see his work is grounded in connection [46]. Bricolage at its core is a network graph, but zoomed out, he captures the fluidity of social interaction. His digital work is similar to Heather Hansen's charcoal sketches [47]. Both bodies illustrate the messiness of interaction, movement, and life. Meanwhile, Guillaume Wiatr literally sketches conversation histories [48]. Though my work engages in drawing in the abstract sense in place of the literal, his work and works by similar practitioners such as those at Drawn to Lead [49] represent how a historical record can tow the line between playful and informative. Bold color and varying stroke weights guide the user's attention.

In *Translations*, I paid careful attention to the visual language. As reviewed in the "Design" chapter, I iterated through several visual forms of representing the conversation data in a manner that expressed meaning and emotion. For this reason, on the spectrum of form and function and technical and artistic, I hoped to situate *Translations* towards form and artistry. The resulting interface should no doubt hold meaning but it should also express empathy and emotion.

2.0.3 Data Humanism & Artistic Representation

In her essay, "Data Humanism, the Revolution will be Visualized," Giorigia Lupi calls for a new era of visualization that leverages art to humanize and personalize data [50]. Since then she has created several art pieces and participatory installations that represent human data as art. Most recently, she released "Unraveling Stories," a series of rugs that visualize endangered rug techniques [51]. In this and similar artworks, Lupi caters the visual language

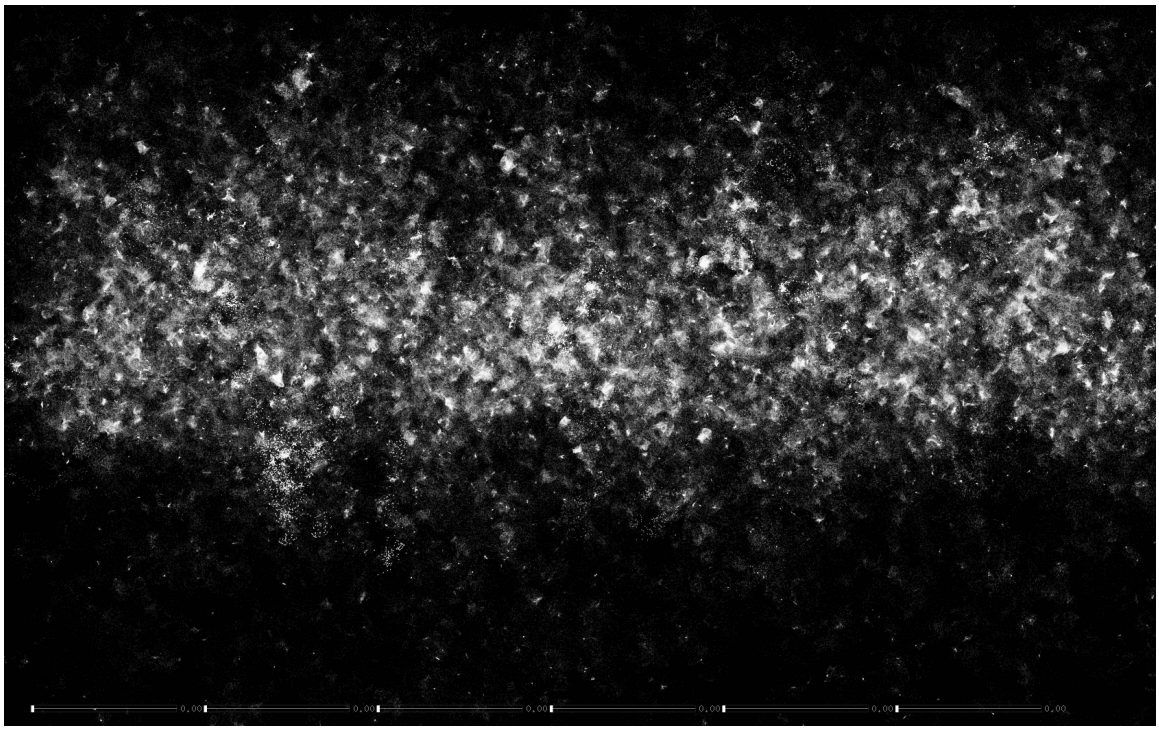

Figure 2-4: Photo of lecture visualization *Bricolage*. Binaura. 2018.

Figure 2-5: Examples of Giorgia Lupi's recent project "Unraveling Stories" in which the history of textiles are symbolically represented on rugs. Photo courtesy of Fast Company, 2023.

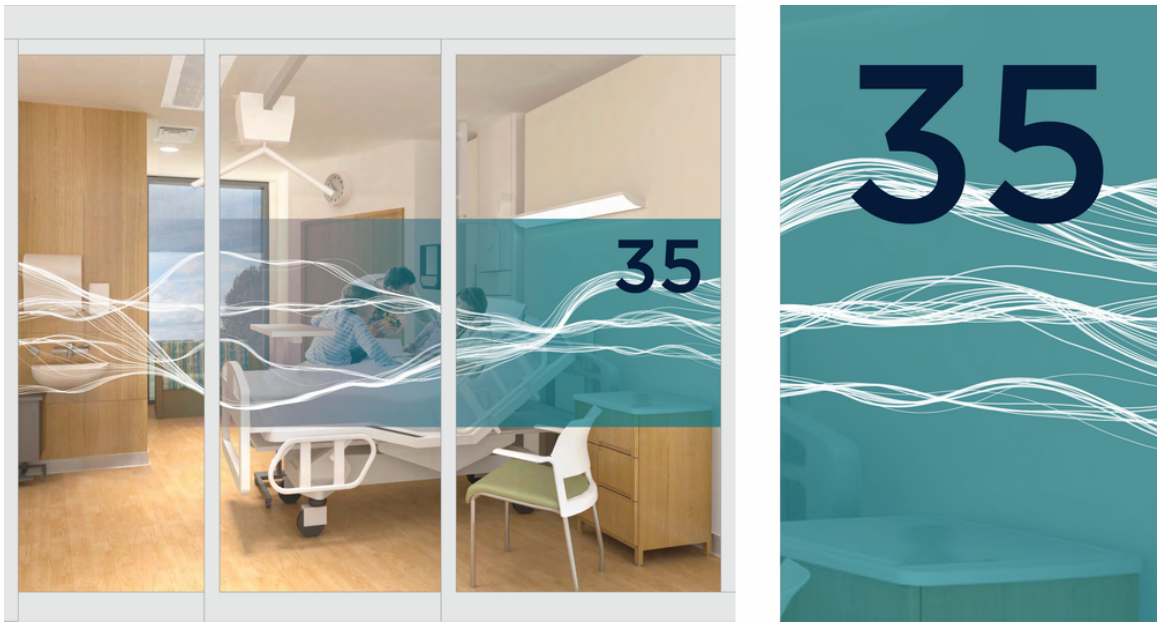

Figure 2-6: A rendering of Posavec's privacy screens for Papworth Hospital. Patient blood-flow to the heart is visualized as waves. Stefanie Posavec, 2019.

to the data and the stories she wishes to tell. The bars and lines of "Unraveling Stories" mirror the looms and patterns the ancestral rug techniques employed. Similarly, Stefanie Posavec employs these principles in her own artistic practice. In her installations for the inpatients ward at the Papworth Hospital, for example, Posavec expresses the patient data in natural forms: waves are stand ins for heartbeats, branching structures for breath, and flows for blood [52]. By leveraging natural and familiar forms, their work not only evokes a sense of balance that is aesthetically pleasing to the eye but also allows audiences to quickly grasp their pieces' meaning and emotional weight. Taking inspiration from this data humanism movement spearheaded by Lupi and Posavec, I believe that art can make the stories and themes present in our corpus engaging and aim to translate the complex themes and data into accessible forms.

For the past year, I have experimented with generative data art as a method of sharing community stories with key stakeholders and the larger public. My intention with each piece in this series has been to give audiences opportunities to truly listen, to linger and deeply engage with the personal experiences shared, and to hopefully see themselves, their

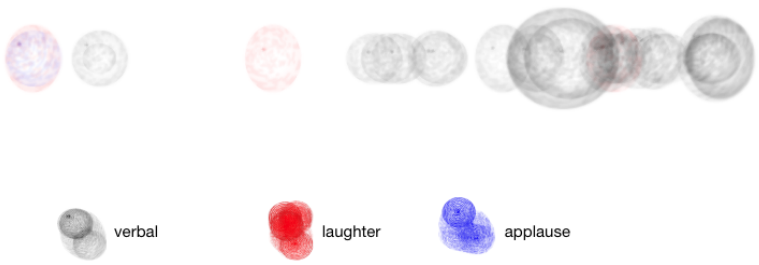

Figure 2-7: Conversation Prints. The utterances RTFC Conversation 36 visualized as fingerprints. Naana Obeng-Marnu, 2022.

values, and their concerns reflected in the collection. I also desire that members of these communities feel heard and seen and begin to feel a sense of agency over and recognize the power of their own voice.

Conversation Prints

When translating personal narratives, accurately representing the voices of the conversation participants is the objective but also the main challenge. Visualization can be a powerful tool to capture complex datasets, but one must be careful to avoid sterilizing the raw experiences of the subjects [50]. "Conversation Prints" was my first experiment in visualizing conversation data. In this piece, the utterances of a single conversation were represented as fingerprints. The result was a timeline of procedurally generated prints of varying sizes and colors coded to resemble those that would be left by an artist working with charcoal.

In his 2015 essay, "Laughter & Silence: The Spiritual Odd Couple," Palmer presents laughter and silence as correlated with sacred spaces [53]. He describes moments when laughter or

silence had facilitated connection to others in his life and muses that these are entry points to meaning and connection. Taking a cue from his writings, I sought out moments of laughter and silence. I listened to a handful of Real Talk For Change conversations were held in person and noted each time someone laughed or the room fell silent and to what in was in response; however, in doing so, I began to notice that, due to the structure of our facilitation and conversation guides, other cues such as "mm" or "yes!" were more frequent in our library and corresponded better to engaging personal stories. I then annotated these and other sounds–applause, pen-fidgeting (signs of nervousness, discomfort), etc–as well. Social mimicry, the often nonconscious act of mimicking and mirroring the verbal expressions and bodily movements of others, mediate prosocial, intimate interaction [54]. So, visualizing such behaviors would also illustrate impactful moments.

The goal was to capture vocal synchrony: how a collective blends together as their individual stories converge. The blending and dragging of fingerprints as inspired by Roth's Multi-Touch pieces became the new visual language for the project [55]. Using p5.js, I plotted the series of prints (utterances) from the manually annotated conversation (RTFC36) over time. In the first iteration, the mapping of conversation data to print characteristics were as follows:

- **Position (start_x)**: Time of utterance
- **Color (type_color)**: utterance/sound type (laughter/applause, verbal)
- **Shape (type_shape_h)**: further classification within utterance type ("mm" vs "piggyback")
- **Width (emphasis)**: prosody
- **Opacity (engagement_alpha)**: collective interaction

The mapping of the conversation attributes to print characteristics in this iteration were arbitrary and required more intention. The distinction between an "mm," "aww," or "piggyback" (currently depicted as the shape of the print), for example, is ambiguous, though

each should carry its own weight. And, highlight topic codes and the thematic tags would be an interesting dimension to incorporate.

The big circle towards the end seemingly captured a moment of synchrony but more can be done to make this more visually impactful. Finally, there is currently no visual distinction between the authors of the utterances. Though one could argue that erasing the individual and emphasizing the collective "we" has its own effect, generating distinct prints for each participant could be a subtle method of reinserting the individual contributions.

Furthermore, this was a purely visual experience. In subsequent works, I strove to create audio-visual experiences that not only represented the voices in a collection, but invited others to listen.

Theme Sort

Following this, I looked to speech duration as an alternate proxy for issues of importance to participants. I visualized and sorted each public conversation highlight, selected and thematically tagged by human sensemakers, in the Real Talk for Change Boston collection into their coded themes. The result was a more playful interactive piece that showed how much time was spent talking about certain issues and perhaps illuminated potential gaps in our tagging practices.

As aforementioned, ultimately, the goal of Real Talk is civic participation and action. Our intent is to bridge divides and bolster democracy through constructive dialogue and the sharing of personal narratives. By surfacing the shared values and priorities of the community, we aim to create impact. Thus, beyond the individual conversation level, we must also deliver insights at scale by visualizing collections of conversations in our library. I ask: how can we influence those who interact with the conversation later and move them to action? And, how can we capture stories and oral histories visually for those who cannot or will not listen all the way through? In furtherance of the mission to translate meaning to third party audiences and stakeholders, I began to consider how the cares of the participants are reflected within the conversation data and how to visualize patterns across conversations.

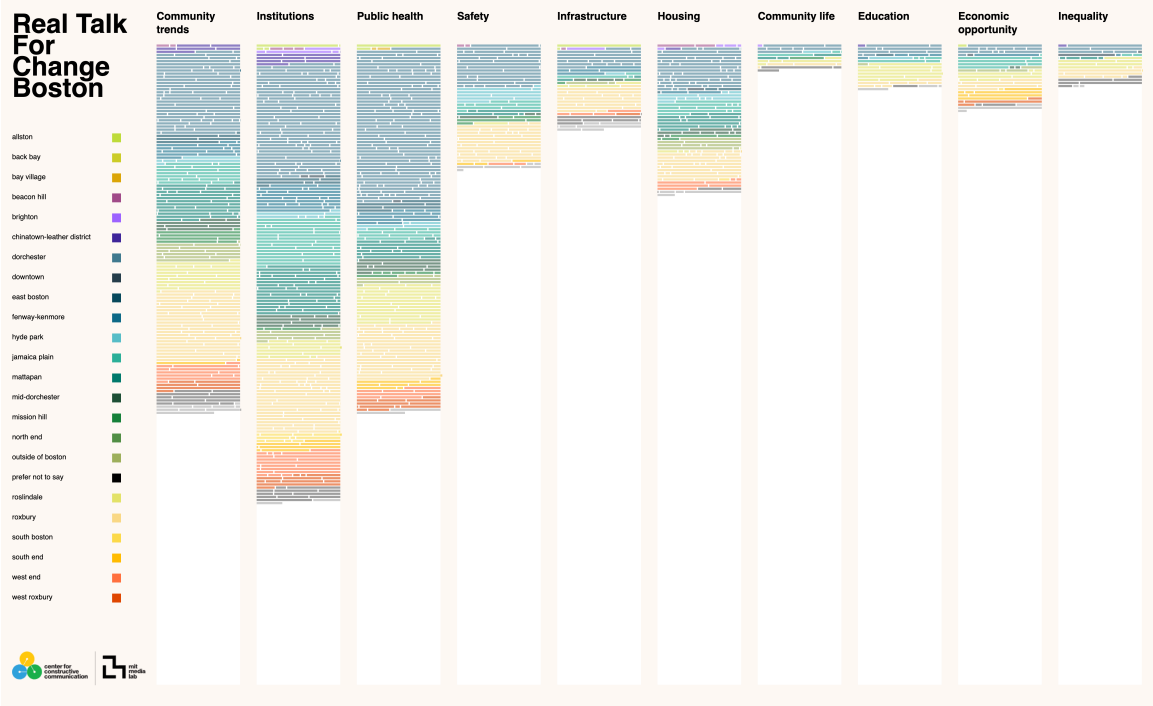

Figure 2-8: Theme Sort visualization. Highlights from the RTFC Collection are sorted into their respective thematic categories. the length of each colored bar is duration, giving you a sense of how much time is spent discussing each topic. RTFC Boston Theme Sort. Naana Obeng-Marnu, 2023.

One such entry point into the desires of the conversants is the sensemaking process during which thematic topic codes are developed alongside community leaders and then applied to highlights from the conversations. These codes can serve as a reflection of values. For the RealTalk for Change Boston conversation set, themes such as "Infrastructure," "Housing," and "Safety" arise. One could simply visualize the frequency with which each of these topics are mentioned; however, is this alone a reflection of the priorities of the participants? How can we begin to re-embed context? Time spent talking about each theme, presented through the duration of each highlight, could be offered as a method to reflect emphasis. Meanwhile demographic data about the participants can also be superimposed.

Through color-coded, stacked, right-justified rectangles that mirror text, Ben Fry's "On the Origin of Species: The Preservation of Favoured Traces" is able to communicate the length of contributions from various edits of his subject [56]. Similarly, Sarva and Nicoletti emphasize bias in the reporting on women through distributed columns [57]. Equipped with these visual metaphors, I set out to visually represent the values the Real Talk for Change Boston (RTFC) participants cared about as indicated by how much time was spent talking about each theme. A column was generated for key themes and conversation highlights, represented by a series of lines, were distributed across each. The length of the line lines were determined by the duration of the highlight, while their color represented demographic information. In figure 2-8, color is neighborhood. At a glance you are able to see that Community Trends and Institutions were top of mind for most participants, while Economic Opportunity was a less pressing topic of conversation. CCC's philosophy with such visualizations is that those interacting with them should always be one click away from the voice as we strongly believe that voice facilitates connection to the story. For this reason, users can also hover over each highlight to hear the speaker's voice.

This visualization began to layer in more complexity into the experience, illustrating more information about the conversation participants and their desires. However, I became skeptical if speech duration was a reliable indicator of consensus. One speaker could easily skew the visualization. Additionally, the theme sort had the same problem as most portal visualizations: scale. The quantity of highlights visualized is so large that it discourages the

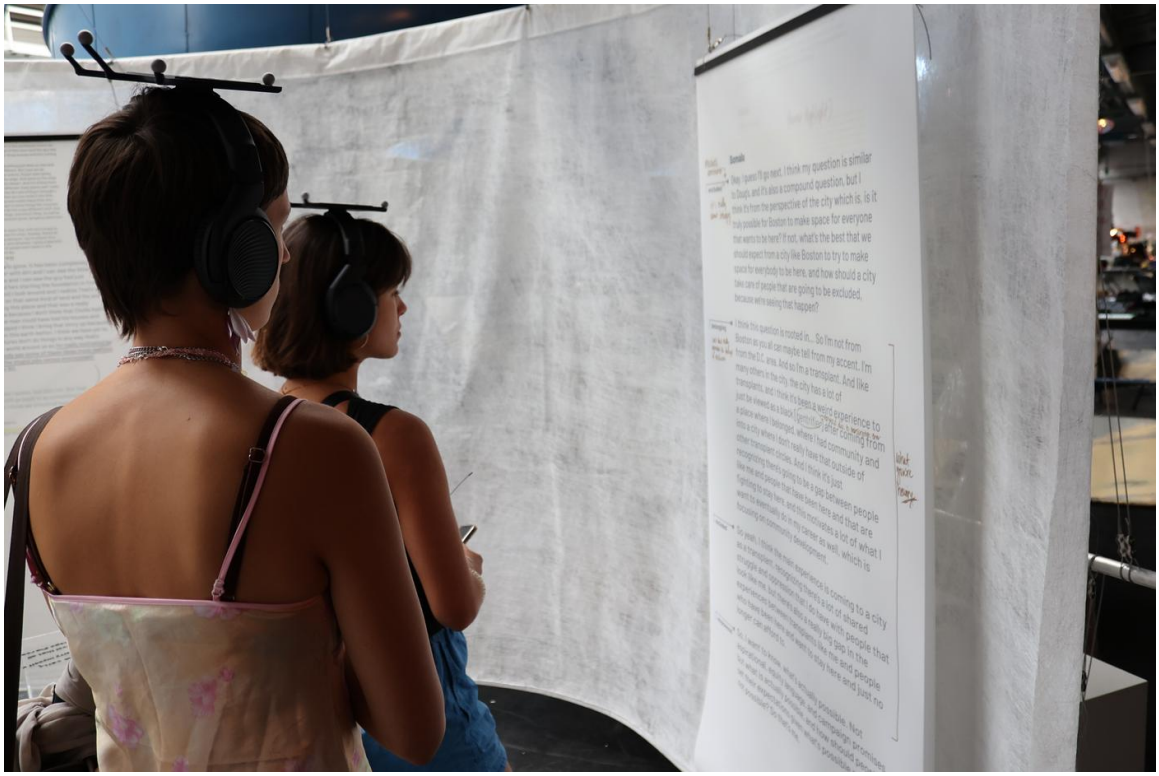

Figure 2-9: Two Ars Electronica Festival attendees listening to the voices in *Collective Echos* while simultaneously reading the transcript banners. Photo by Heather Pierce, 2023.

listener from listening.

Collective Echos

Next, I collaborated with Maggie Hughes, Cassandra Lee, and Spencer Russell to design *Collective Echos*, a spatial audio experience that gave audiences the opportunity to walk through the voices in our conversation library. This new modality allowed people to explore our collections in a new way. Finally, with this physical installation, participants slowed down in the space and spent time with the voices. The stories chosen had a powerful effect on those who entered the space. To provide additional context, I designed banners that displayed annotated transcripts showing the featured excerpts within their original conversations. *Collective Echos* represented the ideal listening experience; however, it raised questions about how to curate and organize the voices within the space.

Figure 2-10: Listener interacting with "Stories in the Rain." The listener is close to the screen, so the volume of the selected highlight is at its maximum level. Photo by Naana Obeng-Marnu, 2023.

Stories in the Rain

"Stories in the Rain" is an interactive generative art piece that explores the effect of embodied movement on deep listening. Stories from the Real Talk For Change Boston conversation collection are visualized on a portrait digital display as raindrops against a black background. An ultrasonic sensor captures the proximity of the listener from the screen. The distance between the listener and the screen is inversely correlated with the volume and focus of selected voices in a p5.js sketch. As the listener approaches the screen, the selected story becomes louder and its nearest neighbors are pushed away from it, creating a void that further draws the listener's attention to the chosen story. And, as the listener stands close to the screen, they can subtly see their own reflection in the rain.

I chose proximity as the mediator of the interaction because I wanted the listener to move–to

buy into the experience. I had hoped the experience would also mimic the natural inclination to lean in towards or approach a person when they are telling you a story you want to hear. With *Collective Echos*, I found that, when we placed audio stories in space and allowed participants to navigate the scene and search for voices with only their ears, our participants lingered [58]. They sat with the stories for a while. Embodied movement added effort to listening, a connection to the stories that is not always found when scrolling on a screen. Now, though they were still a curated set, these were no longer just stories fed to them, these were stories they themselves had sought out. Our participants shared how they preferred this physical form of navigation to traditional digital interactions. Cardiff and Miller's *Osaka Symphony* also enables audiences to control their audio experience through movement [59]. In this installation, audiences can mix audio based on their position in the corridor, curating their own mix. Physical movement affords users a different type of agency, perhaps a deeper connection. I wished to explore this phenomenon further in my piece, though this time applied to a different part of the experience. I did not want the user to click. Movement is not mapped to selection in "Stories in the Rain," but rather volume and focus. The user may not select the voice but they can still decide which they give their attention to.

The main goal of this piece, as with most of my work that engages with this conversation collection, is to encourage deep listening and engagement with the stories that have been shared. The current piece does not achieve that, primarily because of the frequency with which a story is selected. The interval is currently set to 15s, meaning that most stories are cut off and the shorter statements are just replayed in a loop until the 15 seconds are up. This undermines the listening experience.

I specifically did not allow the users to select what they heard because I wanted to test how people would engage with a curated experience. My past work has been choose-your-own-adventure style, and I was recently challenged to try removing user agency, letting the user sit back and listen. Removing choice requires thoughtful curation, it requires a narrative. The random selection here creates some confusion and not a useful confusion.

Additionally, I used a dataset I already had on hand – one that I had cleaned for a visualization where listening was not the main goal ("Theme Sort"). This meant that some of

the audio clips were not suitable for this experience. Some clips were in other languages and others were hard to hear. Some contained irrelevant information or were too short, not containing enough words to convey meaning out of context. If I were to restart, I would have placed more effort on the audio curation and design, a crucial component of this piece. For example, the void that appears around selected drops does not translate as well without the intended effect of a murmur of voices slowly fading to silence.

While I do not believe I achieved my overall goal with this iteration, there were some successes. The piece may not facilitate deep listening, but I do believe it gave an adequate taste of the library. The listeners may not leave with a full understanding of the values and concerns of the community, but they can still hear community voices. They have the ability to hear the tone of the speakers and the urgency, or lack there of, with which a story was shared. Knowing that people were vulnerable can invoke a reciprocal sense of vulnerability. The experience does mediate this sampling. The headphones allow for a focus on the collection. And from my observations of people interacting with the piece, it does seem like the deliberate motion did cause people to slow down despite not being able to choose the stories. Were the audio selection process fixed, I think people would have lingered longer.

2.0.4 Takeaways and Design Principles

From this related work and theory, I determined the following design principles that I carried with me as I transitioned to the design of *Translations*.

- Center the experience of conversation participants. Surface their emotions and desires, when they come together and when they disagree. Prioritize productive collective interaction.
- All stories must be grounded in context. The source must be clear for authenticity and trust.

Figure 2-11: Photo from Heather Hansen's group exhibition *The Value of a Line* at the Ochi Gallery in Ketchum, Idaho. Heather Hansen. 2013.

- Limit the field of view. As I have learned from "Theme Sort" and prior portals, minimalism draws the eye. Instead of representing a whole collection at once, curate a smaller set to kickstart the user journey.
- Utilize motion animation to slow down the user.
- Avoid traumatic experiences. The interface should facilitate calm interactions. Signal to the user potentially triggering interactions.

Outside of this work, I am also an artist. I work with charcoal and ink and pencil as well as digital forms that mimic these modalities. These mediums reflect the values I wish to convey. The limited palettes of my black and white sketches elicit the emotion I seek in *Translations*. These mediums also allow for a smudginess and blending that mirrors the messiness of social interaction. I often aim to replicate these styles in my visualizations

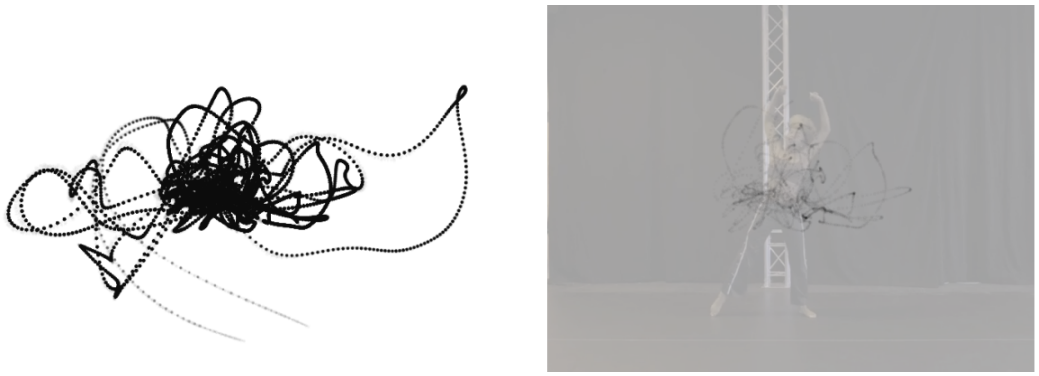

Figure 2-12: Side by side of Inking Movement, a p5.js sketch that turns a dancer's movement into ink drawings. On the left, the ink drawing. On the right, the drawing overlaid on the input video of the dancer. Naana Obeng-Marnu, 2022.

for these reasons and because they provide a calming contrast to harsh digital realities. In *Translations*, I continue to draw on these visual languages and inspirations from print artwork.

Figure 2-13: Invitation Rendez-vous des créateurs. @rdvcreateurs, 2017. Photo Courtesy of Contemporary Type.

Chapter 3

Underlying Data and Systematic Curation

Within the Fora ecosystem exists a library of conversations of varying topics and objectives. It provides a rich set of conversation data to be explored and drawn. Two such collections in this library are the Real Talk for Change Boston (RTFC), a semi-private community collection with selected public highlights, and NYC Community Voices: COVID-19 & Beyond (NYCPHC), a fully public collection. As previously mentioned, the RTFC collection is one I know well. I have experimented with expressing the data of this collection in various visual and interactive forms. It is a collection with which I am now quite comfortable and acquainted. As such, it is with this collection that I began my thesis explorations. The NYCPHC collection is mostly included here for comparison, though data from this collection was used in the evaluation process as will be explained later.

The RTFC collection consists of conversations primarily held in 2021 leading up to the mayoral election that occurred that year. Each conversation consists of 5-8 people. A majority of these conversations were held over Zoom due to pandemic restrictions. The conversation guide used by facilitators in this collection prompted participants to share questions and concerns for the city of Boston.

Metadata	RTFC	NYCPHC
Conversations	72	22
Word Count	530,542	180,892
Snippets	13,560	6,625
Highlights	1,303	941
Speaker Turns	10,808	4,923

Table 3.1: Comparison of conversation metadata for the Real Talk for Change Boston and NYC Public Health collections.

Conversation data from the various Fora collections can be retrieved from Fora's API [60]. It is from this source that I first downloaded the full conversation data and audio for the RTFC collection. Before proceeding, it is necessary to review two key units in which transcript data from the Fora API can be returned. The first and base unit is a snippet. A snippet is roughly a speaker turn as determined by the transcription service Fora uses. If the speaker turn exceeds a certain number of characters, it will be split into multiple snippets. The snippets across the 72 RTFC conversations total 13,560. The second unit is a highlight: the excerpts selected and coded by human sensemakers. Highlights comprise snippets. Furthermore, a request to the API for a conversation's data will return an array of snippets representing the entirety of the conversation transcript.

Typically, highlights serve as the curated set of stories to include in a portal or other share out of the insights from a Fora collection. However, there exist conversations with no highlights and, based on the objective of the conversation hosts, highlights may not be a reflection of interest. At times, sensemakers are merely asked to highlight relevant moments, i.e., when someone answered the prompt. These moments may include captivating stories or may not. And if someone wished to explore a public collection that had yet to or may never undergo sensemaking, without highlights, they would be presented with the conversation audio in its entirety. So, in the quest of systematic curation of moving moments, representative of a whole collection, highlights can be unreliable.

So, I returned to the base unit: snippets, in part, because (1) I wished to develop a curation method that could supplement manual curation, and (2) I wanted to observe the flow of the

conversation in its entirety. Also, I wanted to observe the flow of the conversation in its entirety. Thus, I retrieved the full set of snippets for the RTFC collection. Then, for each snippet, I assembled the available metadata from the Fora API. Unfortunately, the API only returns audio for the full conversation and highlights. To retrieve the audio necessary for the interface and subsequent analyses, I downloaded the full conversation audio, and, using the start and end timestamps provided for each snippet, I extracted the snippet audio from each of the conversations.

A key piece of the sensemaking process is the derivation of thematic insights. Thematic codes are only applied to highlights, so most of the snippets remain uncoded. Due to the volume of snippets, I did not manually code each. Instead, I used an AI Sensemaking script currently in development at Cortico in partnership with CCC. The version I used leveraged Open AI's ChatGPT to assume the role of a sensemaker and apply the existing human-made codebook to the given text. ChatGPT was prompted to apply a thematic code only if the provided text fit the theme, thus a snippet like "Thank you" would not be labeled with one of the RTFC themes. In the end, the script was able to code 4118 snippets.

3.0.1 Emotion Analysis

At this stage, I had assembled a large quantity of metadata related to the RTFC corpus, including the snippet transcripts, audio, and speaker demographics where available. I then turned to answer the question of how this data could be used to represent and reflect participant experience and inform a method of systematic curation. I chose emotion analysis at first out of curiosity. I wanted to note the emotional shifts within a conversation. I was concerned with the effects people had on each other, whether certain individuals swayed the emotional temperature of the conversation. I also viewed emotional analysis as further context against which the final curation and design would rest. In attempting to capture consensus and discord, emotions would then be a great way of contextualizing a speaker's contribution: did their fellow participants feel similarly about the experience they shared? I set out to assemble the emotion data to further characterize the data set, but in summarizing and aggregating the emotions by snippet and then speaker turn, I began to observe the

stories associated with the various emotions. This analysis not only gave me a general sense of the flow of the conversation but also drew me towards individual stories, suggesting that emotion analysis could play a role in both context and curation.

For the emotion analysis, I used Hume's Expression Measurement API [61]. Hume is a research lab that offers facial and vocal expression measurement tools [62–64]. They provide an API that allows one to process an audio clip's language, prosody, and/or vocal bursts (vocal expressions such as laughter or sobs or exclamations like "yay" or "yuck") for its associated emotional expressions. For a given audio file, the API will return an artifact of up to 53 emotions [1] each with a score of 0 to 1. For example, the audio may be scored 0.2 for Adoration and 0.1 for Joy and 0.5 for Anger, etc. Each audio file will receive 53 or 46 scores, based on the chosen expression modality (Language, Prosody, or Burst) [65]. Given the extensive amount of data, a key task was filtering for the most relevant set of emotions. More context for this process will be detailed in the "Design" chapter, but I will relay the components relevant to curation below.

I was initially drawn to prosody. Part of what is lost in visualizations is the tone of voice. Without listening, you may view what the speaker is talking about but their expression is lost. I hoped to recapture this expression and visualize it, so that the listener could grasp a sense of this emotion before clicking to play the audio clip. I had also hoped that prosody would be the most accurate representation of the participants' intentions as tone can greatly alter the meaning of language. However, upon running the analysis, I noticed that one particular facilitator's contributions were being tagged as "tiredness" or "boredom" or even "romance." I found this odd, especially given the fact that I knew this facilitator and these were not words I would use to describe his facilitation style. I listened to his snippet and, while I could see why the AI would label his gravelly voice that way, these labels indeed did not reflect his tone nor intentions. This facilitator was engaged. Yes, he was calm and collected in order to establish a peaceful environment for sharing, but what Hume labeled as "boredom," I identified as interest. Perhaps training a custom model using our corpus would have addressed this issue, but as this was out of scope for this thesis, I chose to revert

[1] Hume returns 53 emotions for Language analyses and 46 for Prosody and Burst analyses.

to more reliable methods of emotion detection for now. Thus, the remainder of this work relies on a combination of linguistic and vocal burst analysis.

I was curious how the participants' emotions evolved over the course of a conversation and whether or not certain individuals or stories had an effect on the subsequent emotions; however, these patterns were difficult to discern at the snippet level, since one speaker's continuous contribution could be broken up amongst snippets. Thus, I aggregated all the snippets to the speaker turn level, resulting in 6589 turns [2], extracted new audio files and ran the emotion analysis again. I set the granularity of the Hume analysis to "conversational_turn" expecting one artifact per speaker turn; however, in some cases, Hume detected multiple conversational turns within a Fora speaker turn. In these cases, I aggregated the outputs by averaging the scores across each artifact to obtain a single set of scores for each speaker turn.

As previously stated, Hume returns up to 53 emotion scores per input. Not all of these emotions are relevant for my purpose. So, I focused on a set of 16 emotions, partially derived from some of the most prevalent emotions in the collection such as satisfaction and disapproval but also including emotions I felt most relevant. For example, Hume reports several secondary and tertiary emotions such as contemplation and admiration, while these emotions may be of interest to some, these do not reveal much about the participant, the story, or the context within which it was shared. Given the objective of the collection, one could expect that participants would contemplate their responses or may be confused about what was expected of them. The readouts of these emotions may be useful in a research or evaluative context–perhaps if CCC or Cortico wished to evaluate conversation guides or facilitation methods–but in the case of surfacing potential stories of interest to the greater public or highlighting community concerns, these emotions may be less useful. At the same time, there are relatively few speaker turns with a Joy score greater than 0.5 [3]; however, I feel it is important to highlight moments of joy in the corpus alongside negative emotions.

[2] At this point in time the Fora Desktop and API displayed and returned only 45 conversations for the RTFC collection instead of 72, resulting in fewer speaker turns. The database has since updated to reflect the entirety of the corpus.

[3] A pilot study evaluating Hume's analysis of the Fora corpuses revealed that 0.5 was the threshold for emotions detectable by humans. See the "Evaluation" chapter for more details.

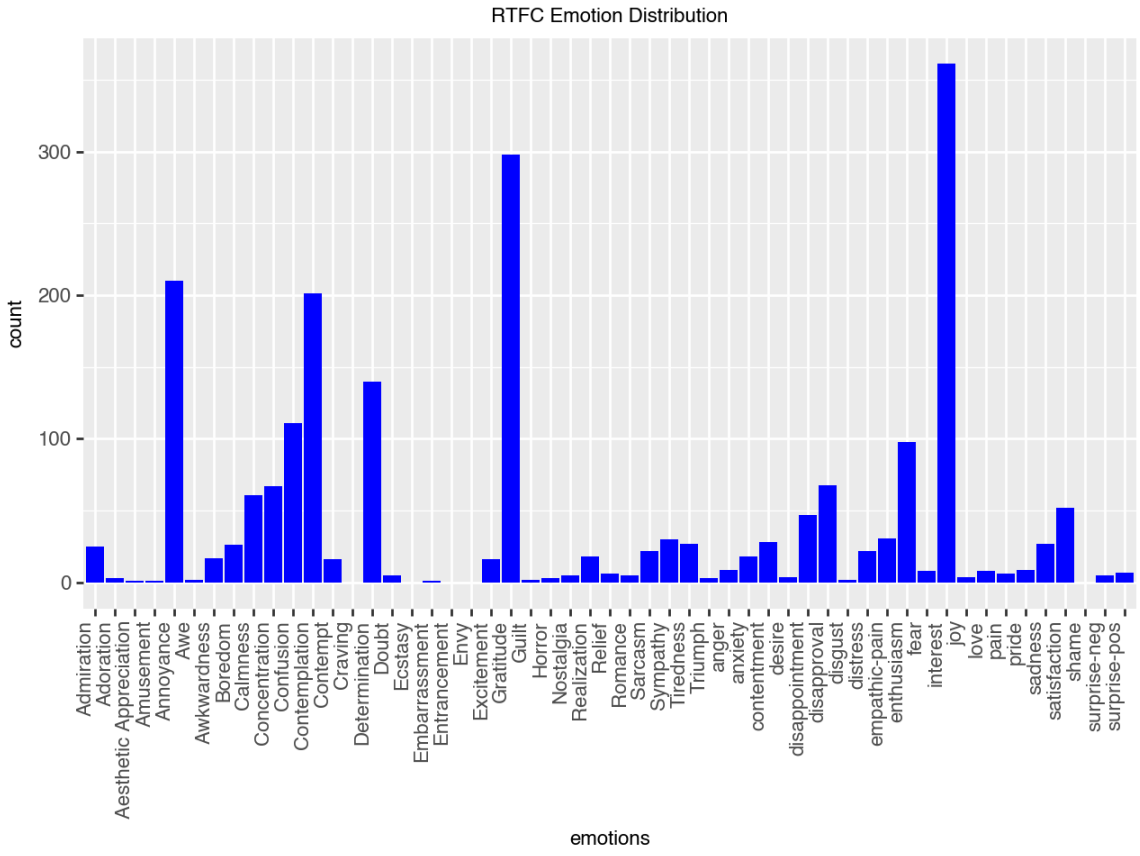

Figure 3-1: Emotion counts for emotions detected with a score greater than 0.5 within the RTFC collection.

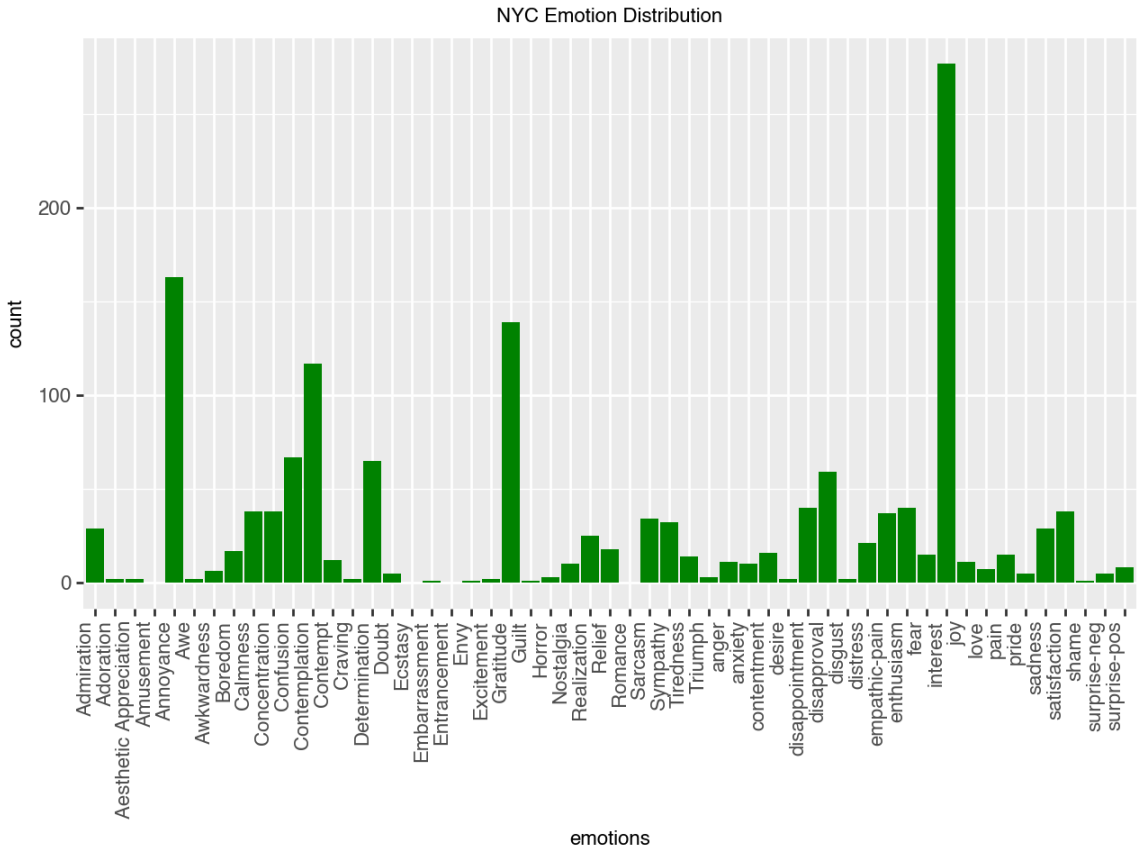

Figure 3-2: Emotion counts for emotions detected with a score greater than 0.5 within the NYCPHC collection.

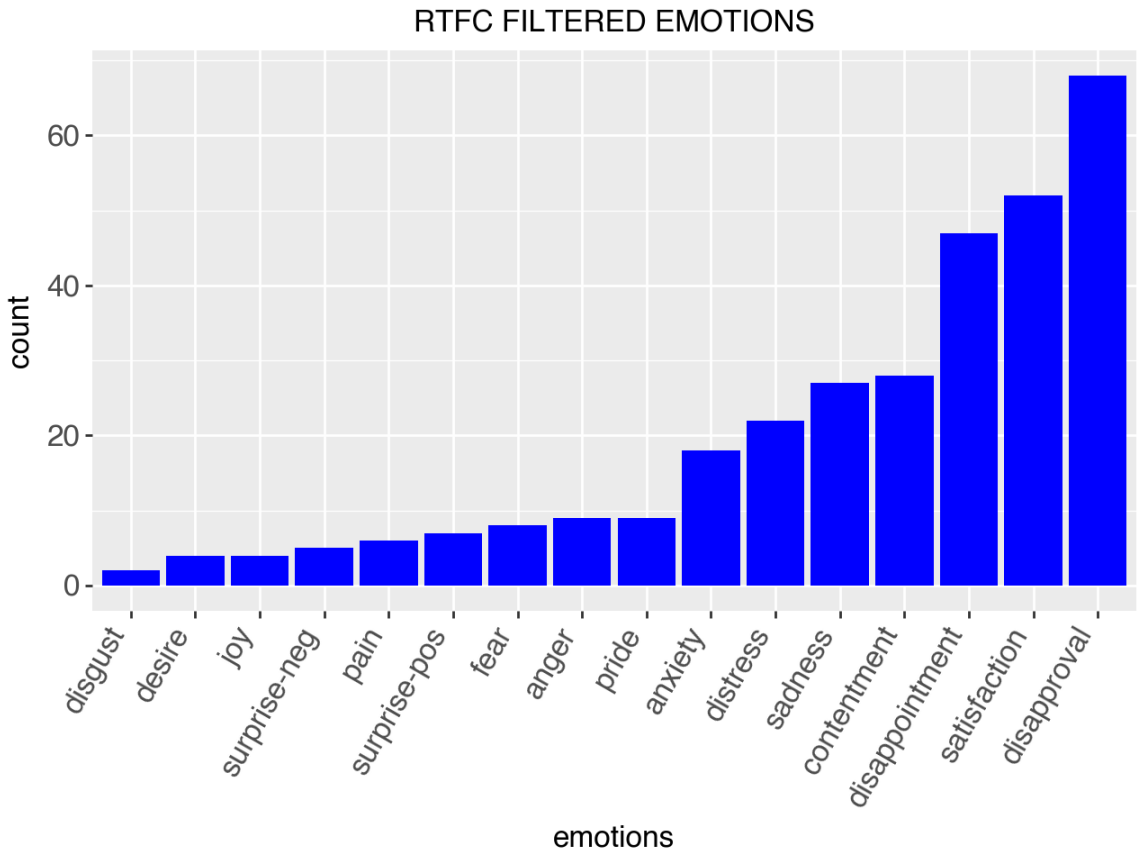

Figure 3-3: Emotion counts for filtered emotions detected with a score greater than 0.5 within the RTFC collection.

To exclude happy moments even if they are rare would likewise misrepresent the participant experience.

Interest and enthusiasm were also included in my initial filtered set. From an evaluation perspective, I was curious whether participants were engaged in our conversations. However, so many of the speaker turns were tagged with interest that no useful insight could be gleaned beyond the fact that participants were generally interested and enthused. Thus, interest and enthusiasm represent cases where I filtered out emotions that are too prevalent. The final set of filtered emotions are 'anger', 'anxiety', 'desire', 'distress', 'pain', 'fear', 'surprise-neg', 'disgust', 'joy', 'surprise-pos', 'contentment', 'sadness', 'disappointment', 'pride', 'satisfaction', 'disapproval.'

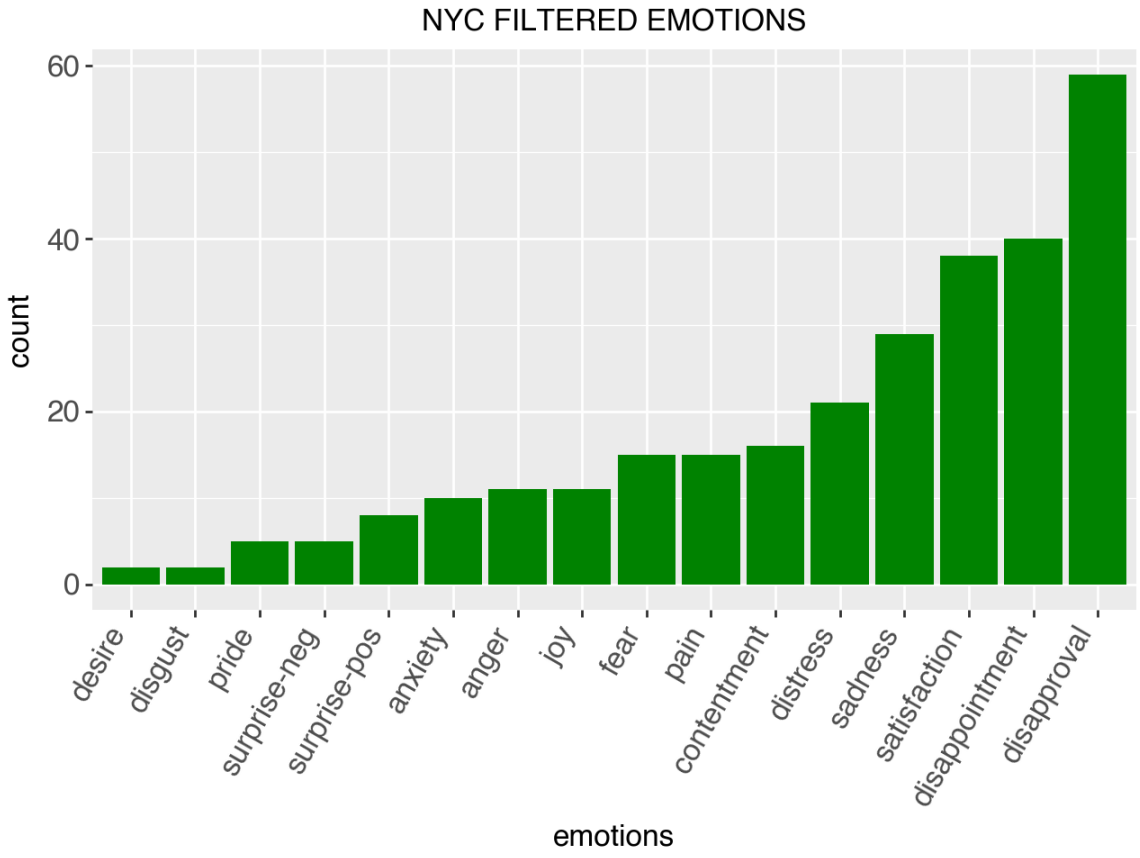

Figure 3-4: Emotion counts for filtered emotions detected with a score greater than 0.5 within the NYCPHC collection.

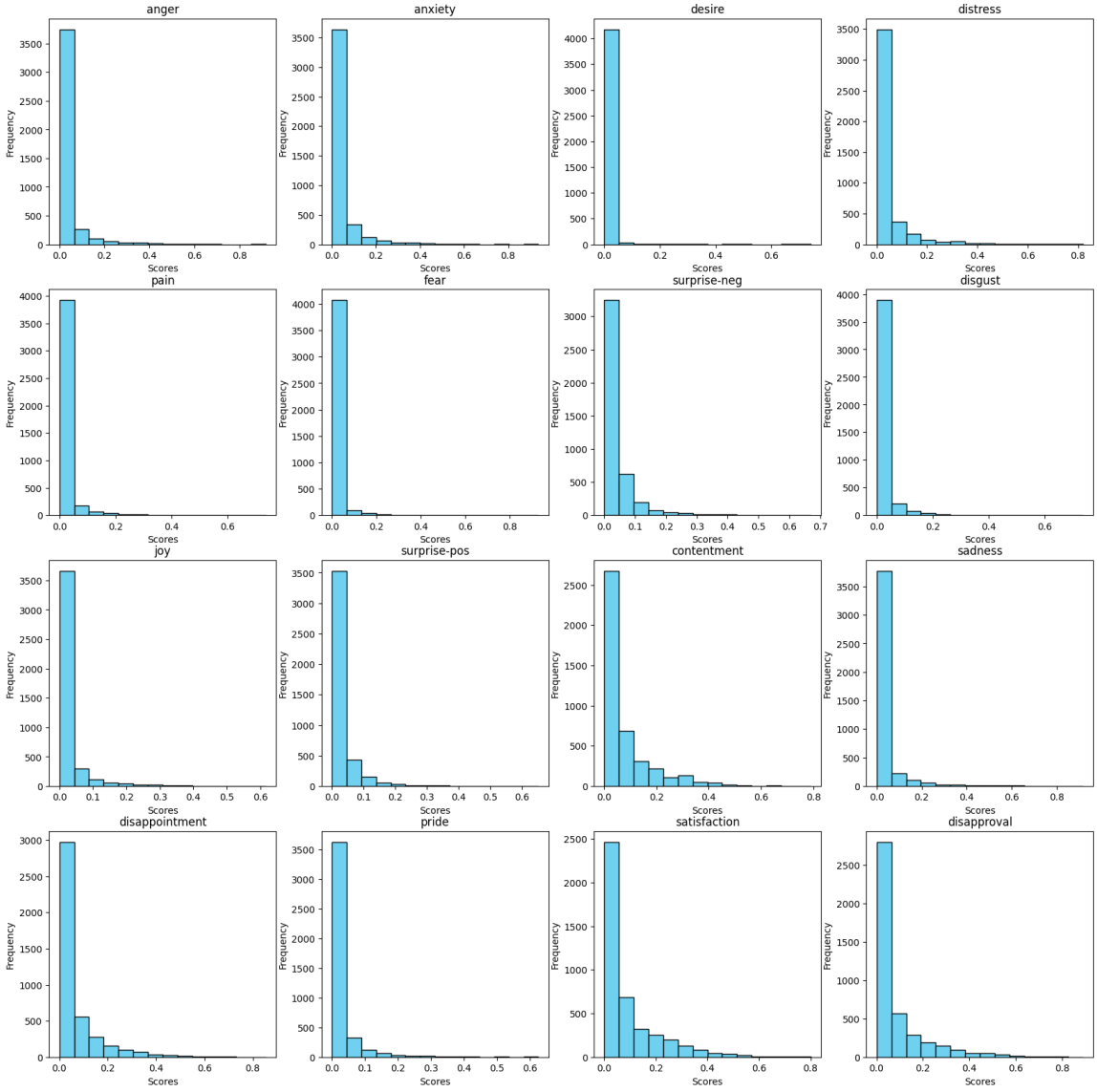

Figure 3-5: Histograms depicting the score distributions for each of the 16 emotions within the filtered set for the RTFC collection. For each emotion, a majority of the scores are below detectable levels.

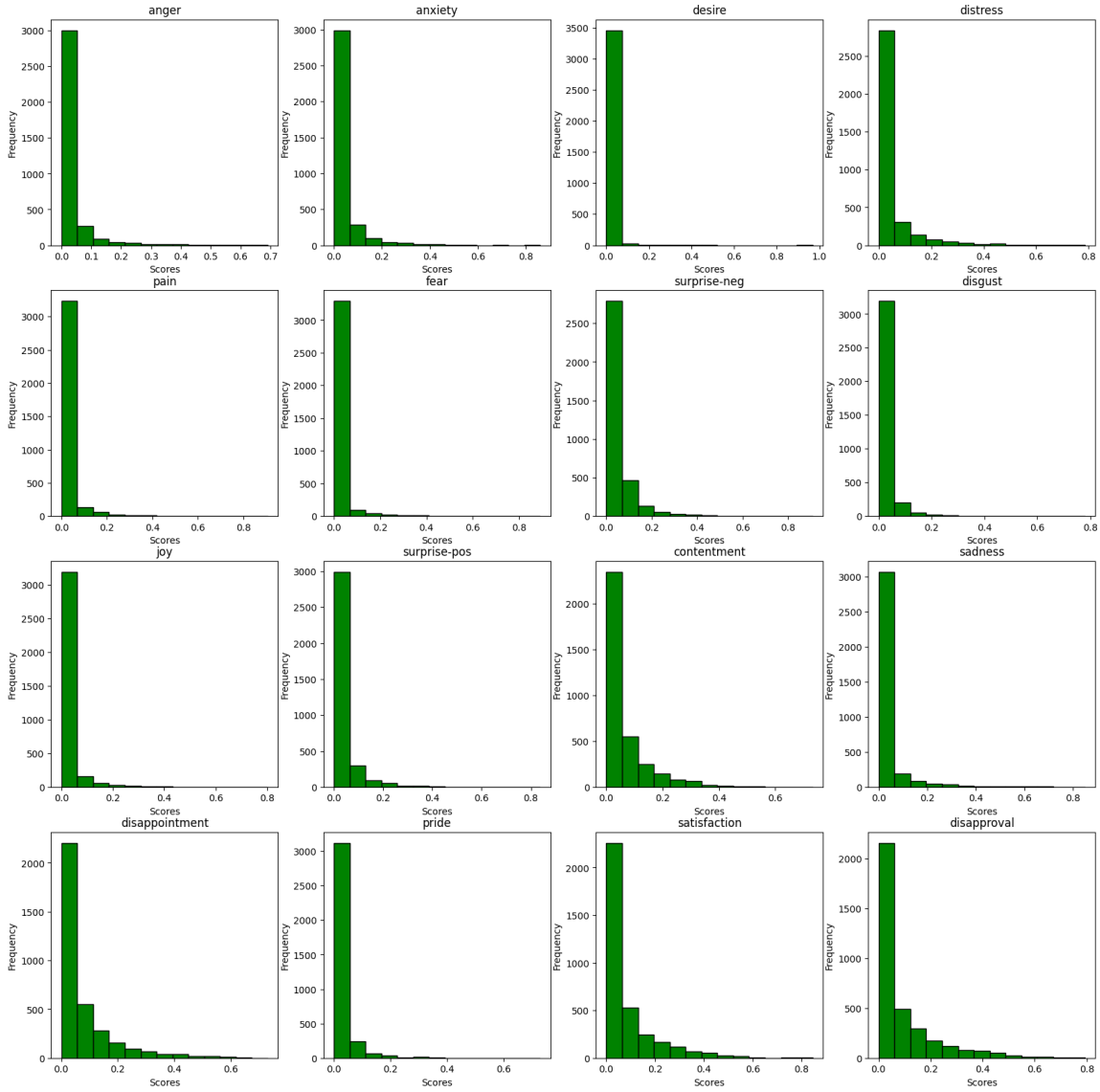

Figure 3-6: Histograms depicting the score distributions for each of the 16 emotions within the filtered set for the NYCPHC collection. For each emotion, a majority of the scores are below detectable levels.

3.0.2 Systematic Curation

Once I had assembled this metadata, I set out to determine how to identify moments of meaning within the collection. I sought out Graham Jones, a linguistic anthropologist and Professor of Anthropology at MIT, who suggested I seek out moments within the conversations where participants signalled cues of reparative listening, potentially discursive markers of surprise [66–68]. Indeed, Chartrand and Bargh and others, including even Stauffer, describe the human inclination to indicate social connection through mimicry [36, 54]. If I wished to elicit a certain reaction from listeners it stands to reason that identifying those reactions within the corpus and presenting those to listeners would elicit similar responses. It was then that I began to develop a method of curation based on vocal bursts, which, I hypothesized, were not only indicators of the speaker of the burst's engagement and participant's interest in the preceding speaker turn but also predictors of the preceding turn's effect on third party listeners. If a speaker turn elicited a burst within the conversation, it was likely to elicit one from a listener outside the conversation. And, if we are to assume that people utter vocal bursts when they are moved [66], the preceding turn is likely one that is moving.

So, to curate the set of moving turns, I would then surface not only speaker turns with detectable emotions from my filtered set but those that also preceded bursts. I focused only on bursts and not general speaker responses because I could not yet guarantee whether a subsequent turn was in response to a preceding turn or was itself a new contribution. Though bursts could have a similar issue, bursts (at least those detected by Hume) are generally reactionary in nature, and false positives would be less likely. To this end, I returned to the snippets for a more granular breakdown and retrieved the burst analysis from Hume. I then compared the start time of the snippets that contained bursts against the start and end times of turns. If a burst was detected in a snippet that started after the start time of a turn but before 5 seconds after the turn ended, I surfaced that speaker turn as a "burst antecedent." I will mention, though, that recent work in CCC led by Maggie Hughes and Brandon Roy to detect new contributions and responsivity could someday allow this process to be refined and expanded to consider all participant reactions.

Chapter 4

Design

The primary goal of the design is to reflect participant experience in an expressive manner, center audio and story to incite generative belief, and leverage motion and interaction design to maintain user engagement. I worked through several iterations of designs that accomplished some or all of these goals to varying degrees. Some designs prioritized voice while others emphasized emotion.

4.0.1 Visualizing Context

The goal of the interface is to translate participant experiences in a manner that guides listeners towards truly listening. The final interface should not only visualize information but serve as an entry point into exploration. I wished to use the conversation dynamics and related metadata to facilitate a listening experience so that those with limited context about the conversation collection could navigate excerpts and begin to understand the community represented in the corpus. There were several dynamics and behaviors I considered visualizing such as responsivity and the evolution of ideas within conversations.

There are many ways to translate experience, but I needed to signal key behaviors visually to support the curation. The initial concepts centered around the metaphor of threads. I imagined how individual contributions formed threads that weave together and that the user

could tease apart with their mouse, unraveling patterns and stories and exposing commonalities through the tensions or slack in the threads. Threads could illustrate connection and separations.

I drew inspiration from textile artists as well as the generative art space. I was lulled by patterns of lines and long standing traditions of woven patterns that held stories. I am still attached to these early sketches and hope to one day return to them.

I let go of this concept for two reasons (1) data availability and accessibility and (2) the subsequent misalignment with the data I ultimately chose as my drawing instrument.

4.0.2 Visualizing Emotion

When I envisioned the intersections of threads, I imagined connection based on shared experiences, chains of responsivity, and demographics. Unfortunately, the data required were either unavailable or immature. Demographic data beyond geographic location is not consistently recorded throughout the Fora collections; the responsivity metrics in development by fellow researchers at CCC were not yet in a format that could be visualized in the manner I intended; and, while theme could be used as a proxy for experience, the categories were often too broad and lacked the granularity I hoped to capture. I would soon turn to visualizing emotions after coming across Rohini Vasudevan's visualization of cat emotions. Seeking similar expressiveness in my own work, the next round of designs dipped threads in emotions, creating the effect of variegated yarn.

Sketches in hand, I began to assemble data. My sketches at that point relied on a single emotion per speaker turn; however, as previously reviewed in "Underlying Data and Systematic Curation," Hume outputs scores for up to 53 emotions for each audio input. I turned to Hume's published literature to understand how I might summarize their emotion artifacts into a single emotion, but I was instead convinced of the necessity of complexity in this case. Hume asserts that humans rarely express a single emotion: it is through the interpretation of multiple emotional cues that we make inferences and decisions about how we wish to

Figure 4-1: Mock up of Conversation Threads, an early iteration of *Translations*. Each line represented a conversation and user's could lift related stories by hovering across the lines.

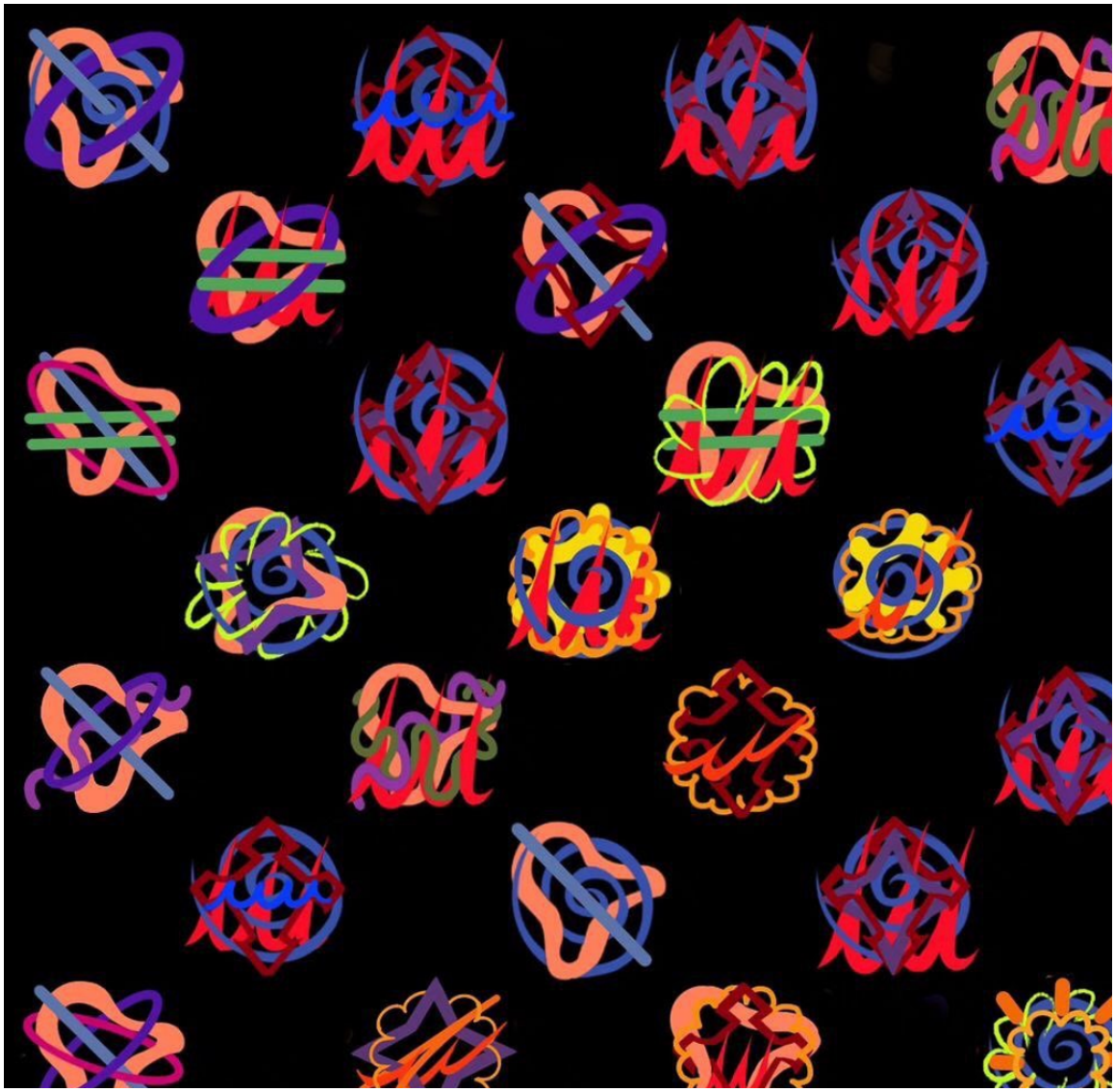

Figure 4-2: Rohini Vasudevan visualization of cat emotions using Hume data. The artist had used Hume to capture the emotions of cats, crafting an expressive and whimsical piece. Courtesy of Data Viz Society.

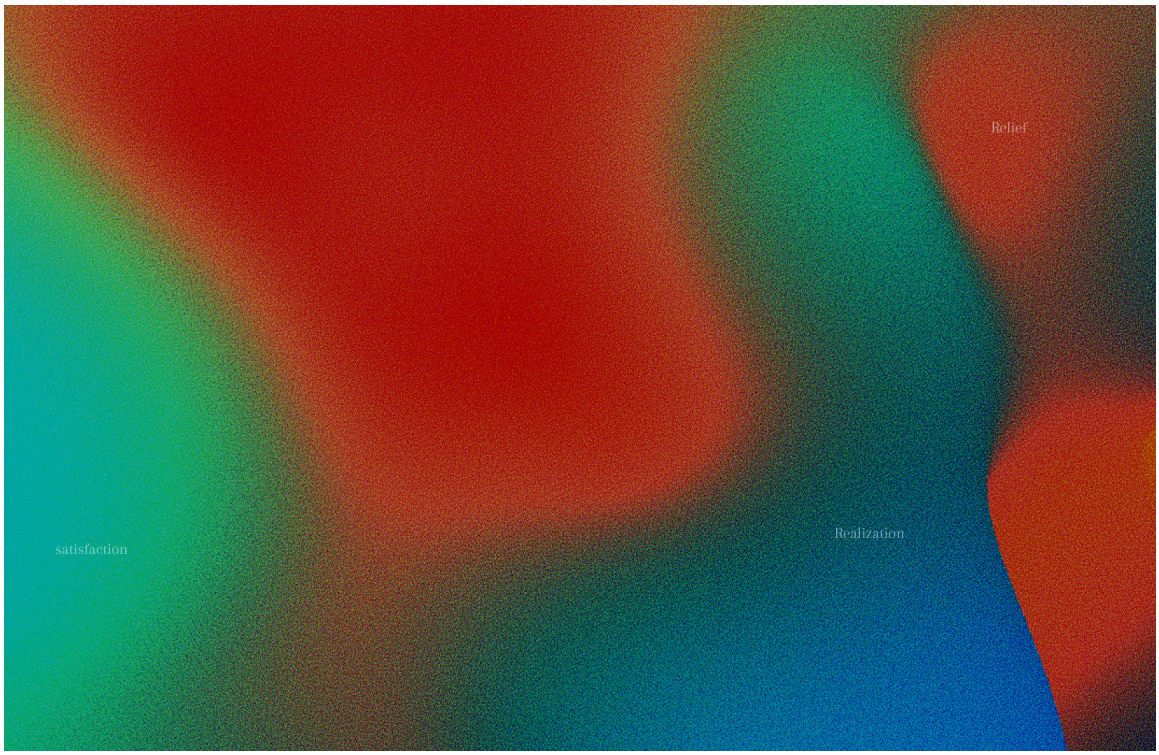

Figure 4-3: In this early prototype, the top five emotions in each speaker turn were displayed as a gradient mesh.

respond [63, 64]. Thus, the next round of sketches aimed to capture the complexity and fluidity of emotions.

I was also inspired by interactive websites like SETAPRINT's which allow the user to interact with fluid, colorful flourishes.

Using React Three Fiber, a three.js library for React, I assembled an early prototype applying the color gradients to each snippet in a conversation. These snippets appeared in a line through which the user could scroll. Ultimately, the user would be able to scrub through the audio associated with each snippet as well. I intended that the user unwrap the orbs they focused on so that the emotions it contained could be fully viewed.

It was at this stage that I began to develop the color key. The first prototype of the orbs assigned color based on the D3.js' Spectral color scheme. Even with the key at your side, it was difficult to discern emotional patterns. Next, I used the colors Hume applies in their

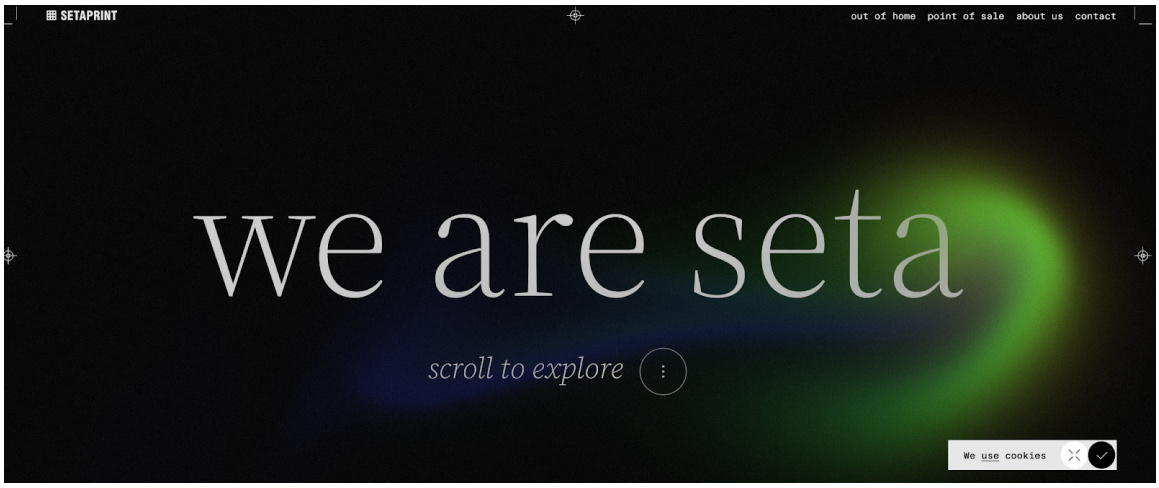

Figure 4-4: Screenshot of the landing page for SETAPRINT, a European print shop.

Figure 4-5: React Three Fiber prototype of snippets represented as orbs. The top 5 emotions of the snippets are applied as a gradient mesh material of each sphere, and the user can scrub through each snippet horizontally.

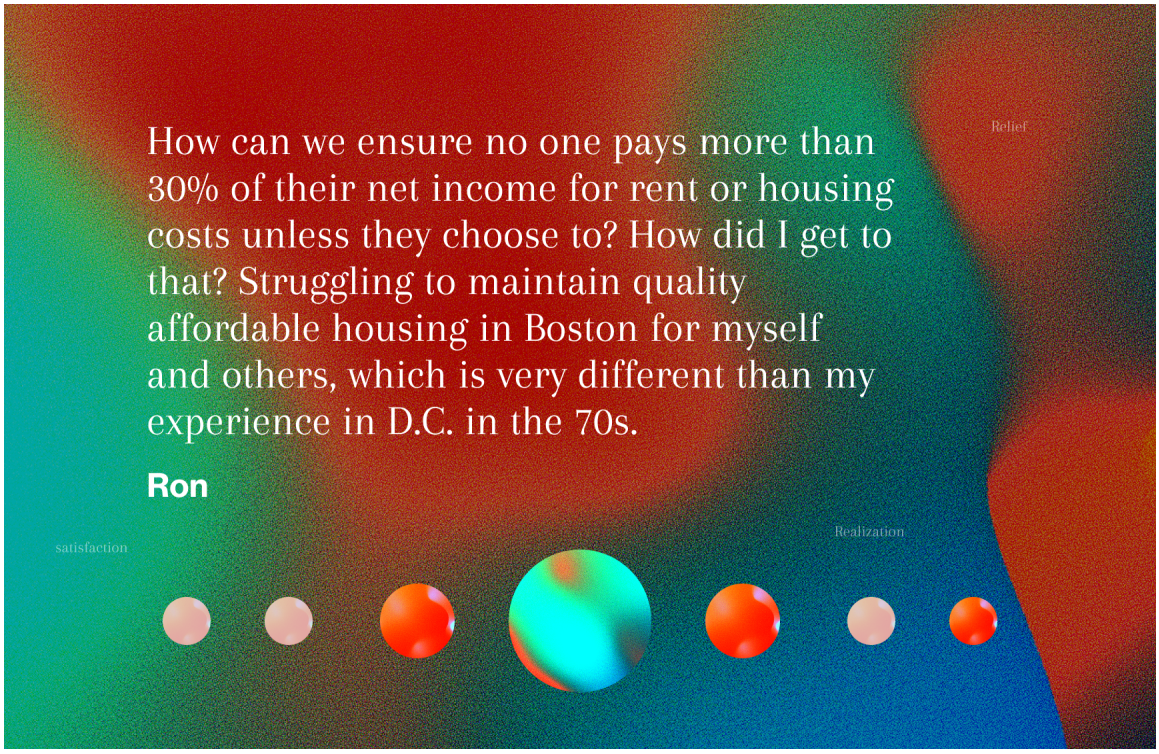

Figure 4-6: Mock up of the *Translations* Orbs - Emotions Gradient interface. When a user focused on a snippet or speaker turn, the background of the page would change to a gradient mesh representing the top 5 emotion colors of that speaker turn. The associated transcript would also display above the timeline.

playground. However, color similarities between seemingly disparate emotions continued to confound the experience.

I then mapped the emotions onto a color wheel, manually assigning emotions to colors based on implicit associations I had. Like most people, I associated anger with red, joy with yellow, and sadness with blue [69]. I associate anger with fire and passion which are generally associated with the color red. Similarly, pain, often related to or even a consequence of fire and anger, I placed in the red color space. Fear and anxiety, which I saw as related to anger and pain but not as inherently negative, I placed in the orange color space. Desire landed in green because I associate desire closely with green-eyed envy. Some emotions came easier to me than others. Pride I had initially placed in the yellow/green space because I saw it as a joyful expression, linked with desire, but its placement did not feel right to me. For emotions such as these that I was unsure of, I consulted literature on emotion and color association [70–73]. These placed pride in the purple spectrum. Though I could not place it initially, it felt right there. Pride is almost regal and royal, adjectives often associated with purple.

I am fond of the sketches and prototypes from this phase as they represent the emotion and context I had hoped to express. Here, Errante's memory pool is made tangible.

Unfortunately, while these sketches captured the expression and reflected the artistry I sought in my work, they were not fully functional. There was too much emotional data present to properly signal to the user what they should listen to. While participant experience was depicted, the story as audio was left behind. Indeed, I struggled to integrate the audio experience beyond the scrubbing behavior previously described.

It is at this point that my advisor directed me back towards generative and plotter art. We were both drawn to a piece by Gunther Kleinert [74].

Simultaneously, I returned to my design value of calmness. I wanted to construct an interface that breathed. I drew inspiration from printwork, particularly Riso prints. The effect of ink absorbed into paper captured the visual experience I wished to create. Bold type against softly drawn lines, ink splotches and splatters against expanses of negative space–these

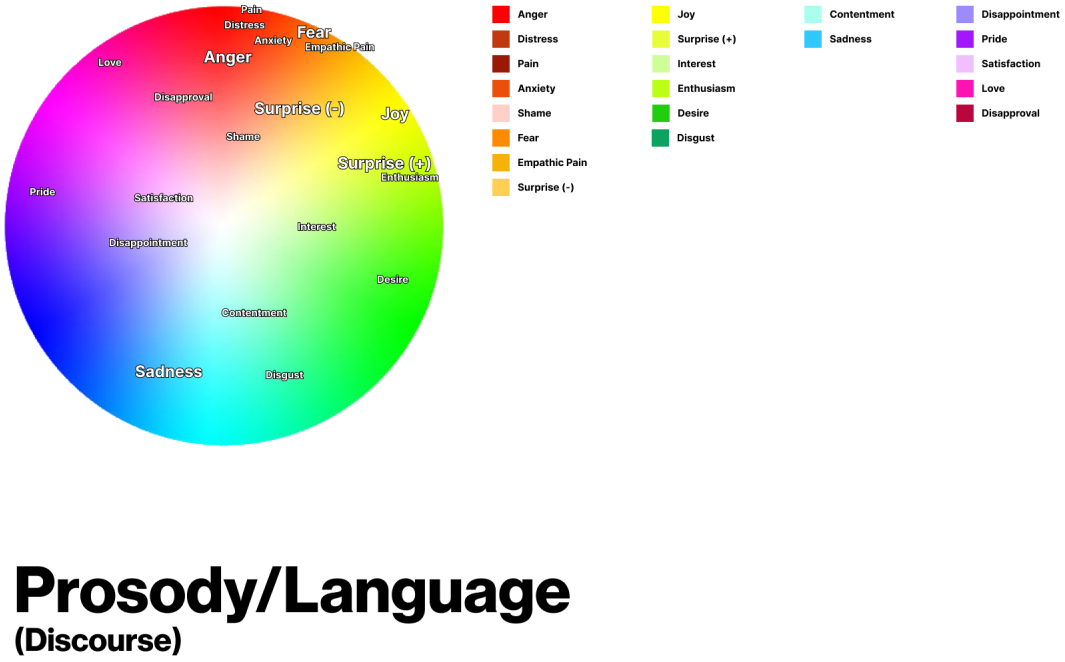

Figure 4-7: My manual mapping of 21 Hume emotions to an RGB color wheel and the associated color key.

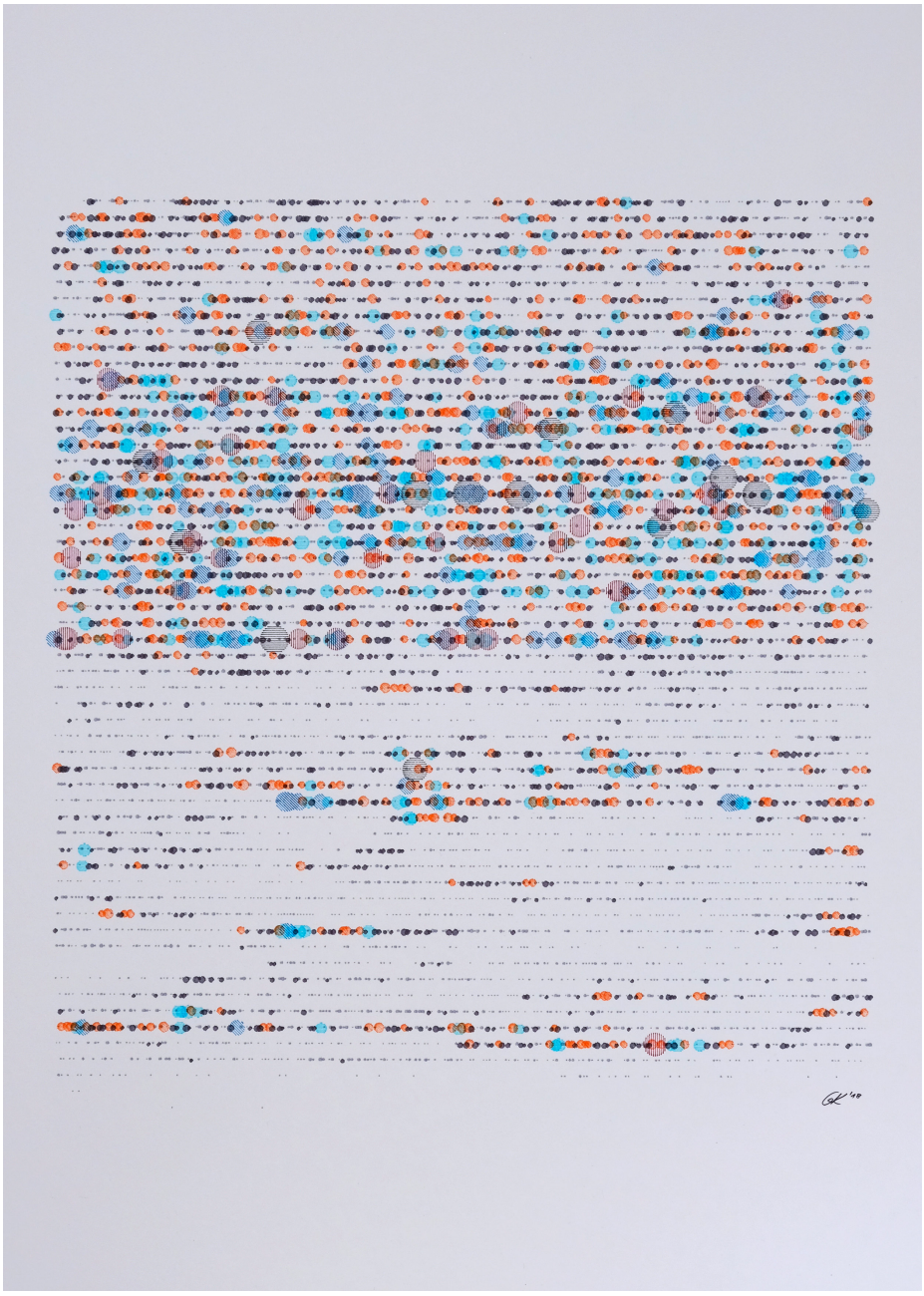

Figure 4-8: “The Köln Concert part II c” Keith Jarrett. As part of the I CAN SEE MUSIC series. Gunther Kleinert, 2018.

relieve the burden on the eye. And, again, the idea of drops on a canvas evoke imagery akin to that of Errante's memory pool metaphor.

Thus, the next iteration of *Translations* was born. I began with quick p5.js sketches to test the concept. These sketches also allowed me to evaluate the curation in context. I began to experience the types of snippets and speaker turns my system exposed.

It was also during this phase that I chose to eliminate interest and enthusiasm as emotions of interest due to their overrepresentation. As you can see in Figure 4-13, where Interest and Enthusiasm are green, their presence made it difficult to discern meaningful patterns in the corpus.

Once I had the basic form, I returned to Figma to finalize the visual language and style.

I developed a prototype in React and presented version one at an event with professionals from various industries. I remained beside the interface to guide users and answer questions. These guests included many unfamiliar with the frameworks and conversation corpuses with which this thesis is concerned, thus they were great examples of the third party public I ultimately wished would engage with my work. Furthermore, given the nature of the event–a research symposium of sorts–many who attended were generally disinterested in artistic work. Thus, the interface piquing any guest's interest would signal success. During this event a handful of professionals interacted with the prototype, though most were only interested in the underlying AI. Due to the nature of the demo and the surrounding environment, the interface was not set up so that people could listen, so I largely tested the interaction. The feedback I received was mostly functional. At this point, there was a lot of noise in the experience. Hovering on a theme or speaker turn revealed too many related turns. So many offered suggestions regarding how I could further curate the experience. I was also encouraged to expand the work to include other languages.

Ultimately, I disabled the interaction where turns that shared themes become colored on hover. The intention here was to surface shared experiences, but with themes alone to represent experience, the interaction oversampled turns and did not reveal more than could be accomplished with thematic and emotion filters.

Figure 4-9: Annotation brush pack example from Bracken Design.

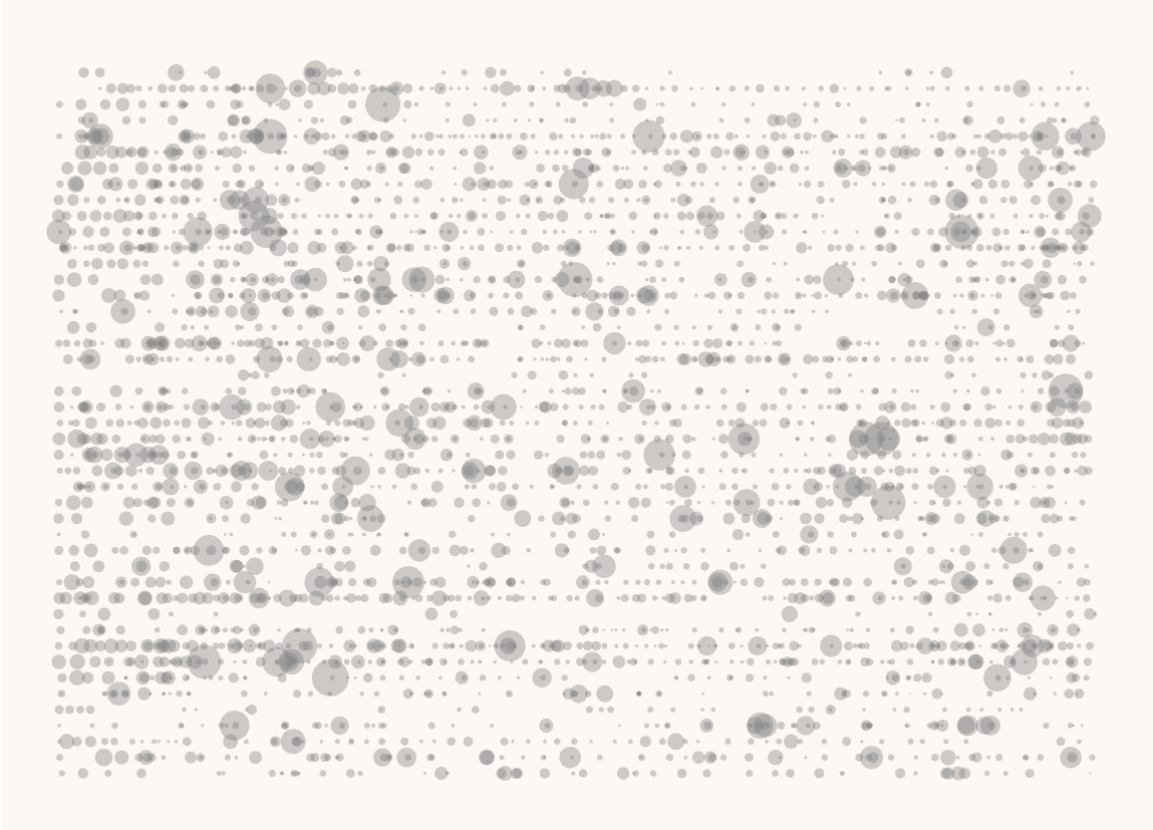

Figure 4-10: Initial plot of snippet emotions for snippets the preceded vocal bursts in grayscale.

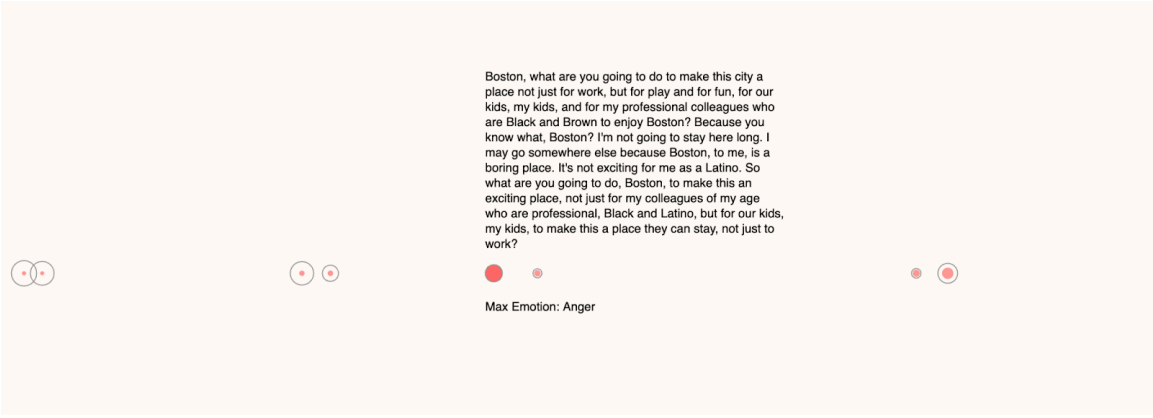

Figure 4-11: Snippet burst analysis prototype filtered for Anger.

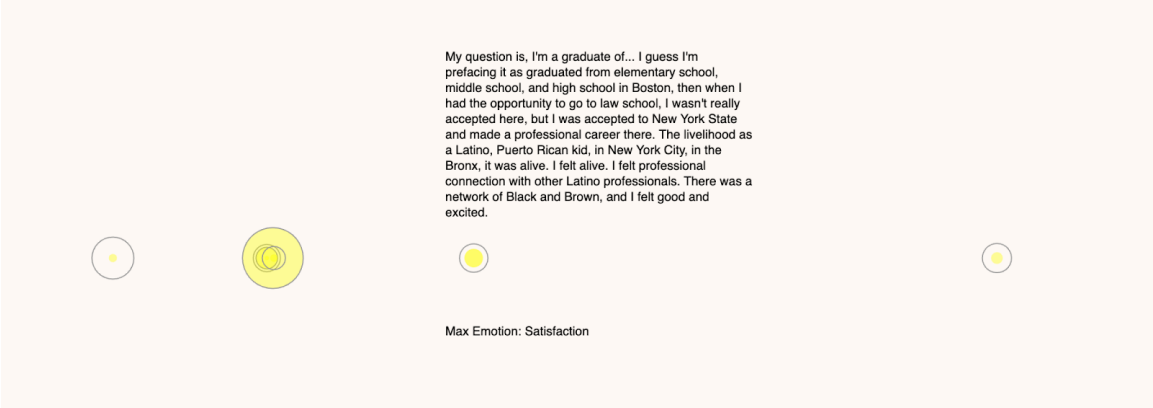

Figure 4-12: Snippet burst analysis prototype filtered for Joy.

Figure 4-13: Initial plot of snippet emotions for snippets the preceded vocal bursts in full color.

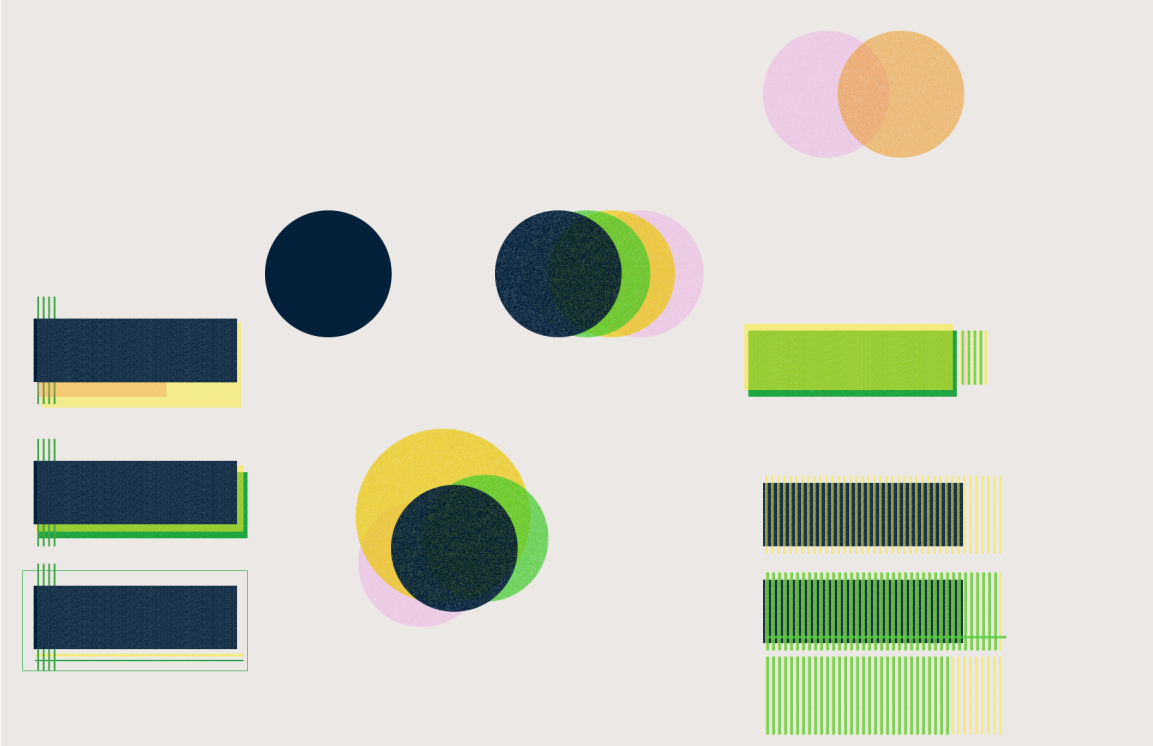

Figure 4-14: Figma sketch of Riso print imitations, testing blending and layering of color in digital space.

Figure 4-15: Figma sketch of the *Translations* interface in Riso Print Style.

The most helpful learning I gathered was in regards to its motion design. In its idle state, this version of the interface included an animation of each speaker turn dropping on the canvas. One by one, they slowly bled into the page. I left this on for some time after the main demo period had ended and observed as many stepped in front of the screen to watch the drops slowly appear. This simple interaction indeed caused people to linger. However, given the quantity of turns, it was too slow and took too long to cycle through all of the turns, so viewers did not get the chance to see the full breadth of the collection.

Following this event, I updated the interface and set it up as a standalone installation in a common area. The space was regularly frequented by a steady flow of students, faculty, and guests–again, third parties likely to have no connection with the corpus or its conversation portal. In the background, the interface collected user interactions. For the first few days, no one interacted with the interface, so I added an animated start screen, after which, several individuals were enticed to test the interface over the course of 7 days.

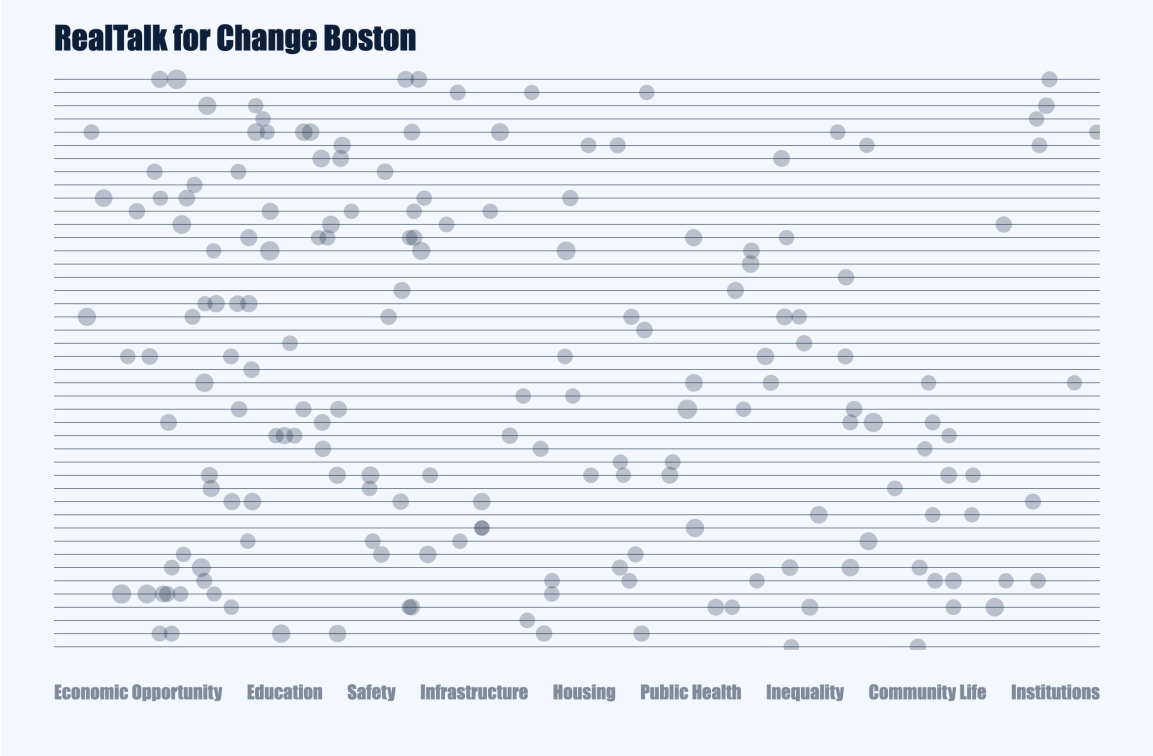

Figure 4-16: Screenshot of the initial prototype of the *Translations* interface. Each line represents a conversation and each dot a speaker turn. Thematic filters can be selected at the bottom of the screen.

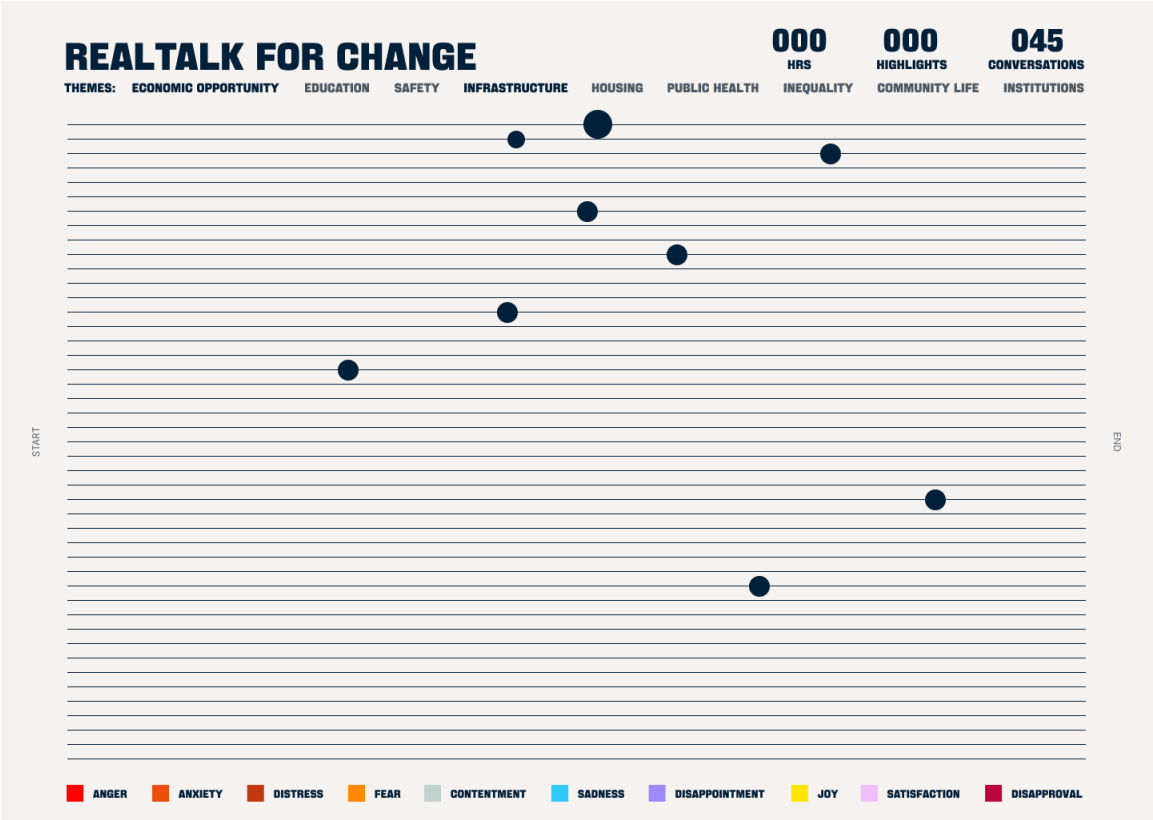

Figure 4-17: Figma Mock Up of the final *Translations* interface. Thematic filters appear at the top while emotion filters appear at the bottom of the screen. Axes labels for the start and end time of the conversations appear to the left and right of the screen. Collection statistics including number of highlights and conversations appear in the upper right corner.

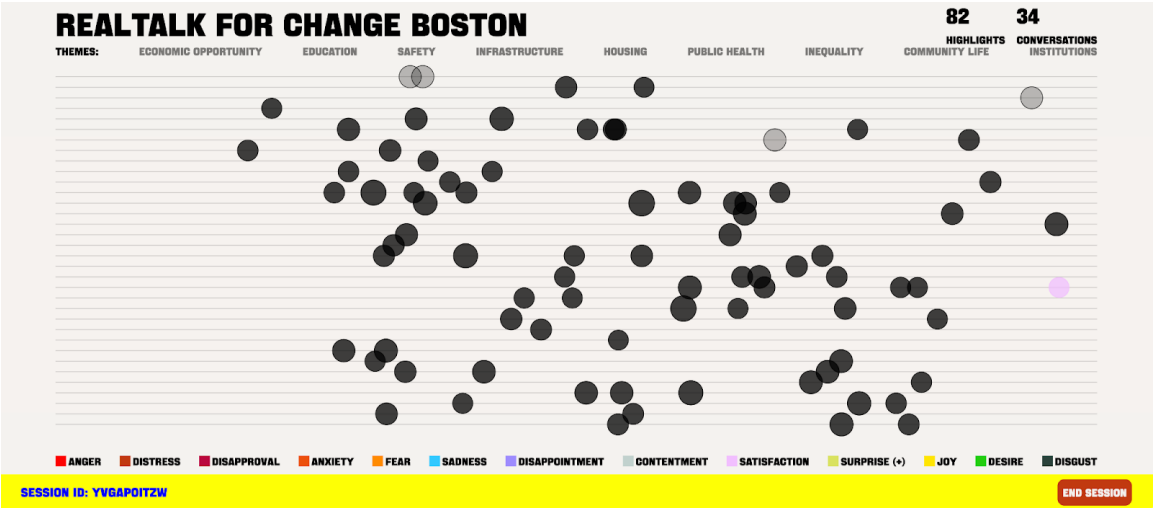

Figure 4-18: Screenshot of the *Translations* alpha test interface with a selected speaker turn. This still represents the end of the animation sequence when the turn audio has finished playing and the remaining curated turns have faded back into view.

In this final state, the interface presents 34 conversations. These are conversations that contain highlights curated using the burst antecedent curation method.

The lines are ordered on the y-axis based on the time in which they were recorded with the earliest conversation at the top and the most recent conversation at the bottom. The x-axis represents time within the conversations, with the start time at the left and the end time on the right. Within each line, curated highlights from those conversations drop and bleed into the page. Their animation incorporates a drying effect with the opacity lightening slightly as the drops settle into the page.

Users can cross-filter themes and emotions. Each drop reacts to the user's mouse on hover. When they click on a drop, all others disappear and emphasis is placed on what they are currently hearing. A ring representing the duration of the drop appears and the colored drop slowly bleeds to fill that outline as the audio plays. By the time the audio is complete, the circle is filled. Finally, the remaining drops return to the page.

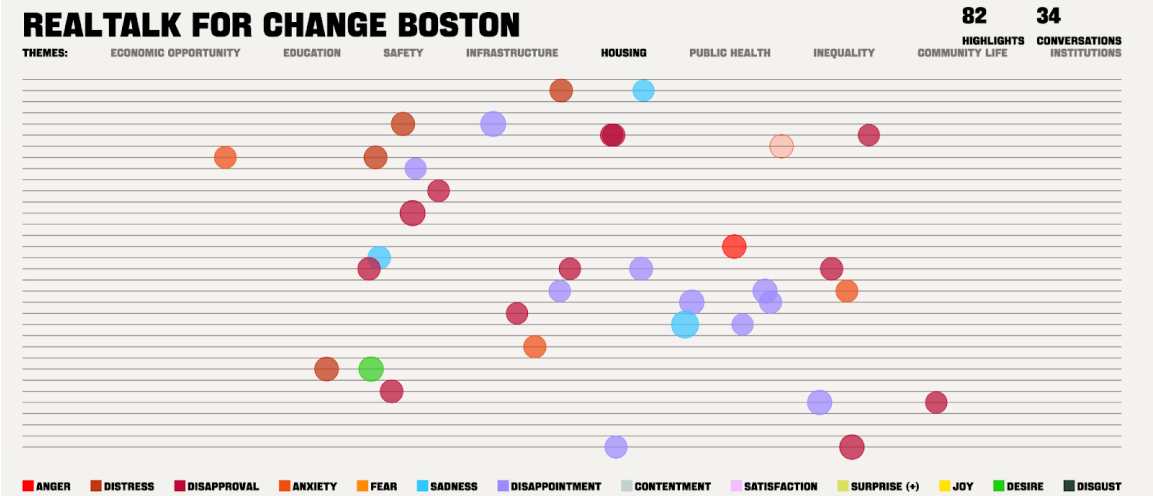

Figure 4-19: Screenshot of the *Translations* alpha test interface with the "Housing" thematic filter selected. Several ink drops representing curated public speaker turns in the RTFC collection match the theme and express a variety of emotions, mostly negative.

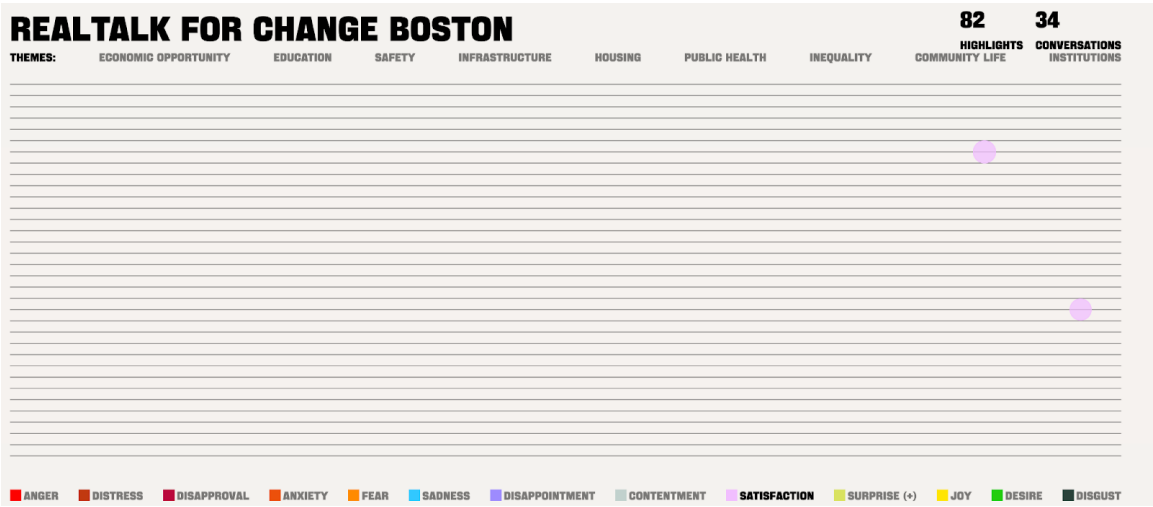

Figure 4-20: Screenshot of the *Translations* alpha test interface with the "Satisfaction" emotion filter selected. Only two curated speaker turns express Satisfaction (above the threshold of 0.5).

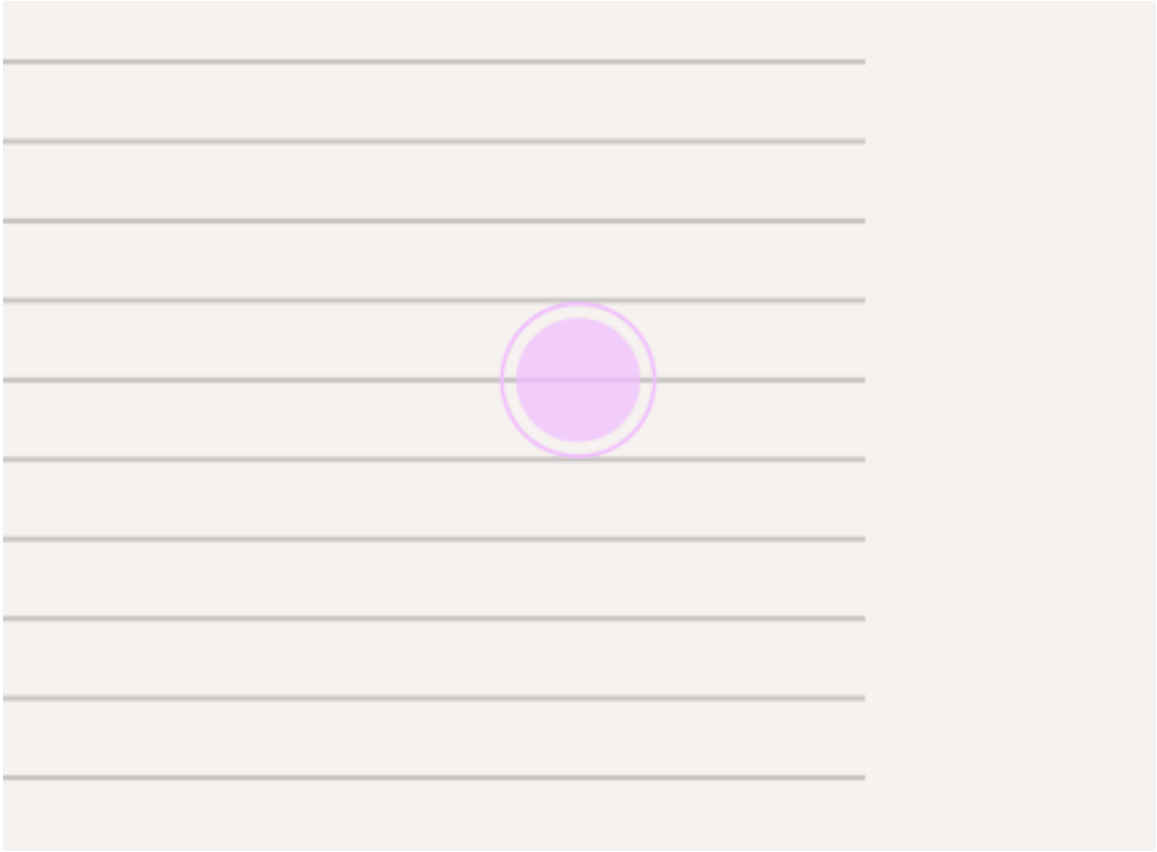

Figure 4-21: Still of selected speaker turn while playing. As the audio plays, the ink drop grows in size to fill the outer stroke which represents the duration of the turn.

Chapter 5

Evaluation

The evaluation of this thesis occurred in three parts. First, I verified the accuracy of Hume's language analysis on a Fora corpus. Next, I sought to confirm whether audio curated according to my burst theory and thresholds had a significant positive effect on empathy compared to a control. Lastly, I set up and shared the final interfaces in several contexts to glean how the general public engaged with it. I tracked user sessions to monitor engagement and listening behavior and conducted interviews to solicit feedback about the experience.

5.0.1 Hume Validation Study

In the first study, I wished to validate that Hume was capable of accurately labeling the motions in our corpus. Separate from, but ultimately in service to, the overarching questions of this thesis, the study aimed to answer the following research questions:

- **RQ 1**: Given the same audio transcript, do the emotion labels applied by humans align with Hume's emotion detection?
- **RQ 2**: What is the minimum emotion score detected by Hume required for humans to recognize a given emotion?

The first line of inquiry was a validation check while the second was for threshold setting. If humans were unable to consistently detect an emotion on a speaker turn with a score below a certain value, I wished to filter these out as presenting these turns as the selected emotion would be inaccurate. To this end, I ran a pilot study using the Prolific platform.

Participants

The study included 25 participants (44% male, 56% female) recruited through Prolific. Only fluent English speakers residing in the United States or United Kingdom were allowed to participate. 84% of the respondents identified as White, 8% as Black, 4% as Asian, and 4% mixed race. Their ages ranged from 22 to 79 ($\mu = 41.4$, $\sigma = 15$ years). Participants were paid $3 for their time. The median time taken was 5 minutes, 5 seconds with an average reward per hour of $34.40.

Prospective participants were also screened out if they failed the attention check question, a sample comprehension question for a nonsense sentence "The blue flamingo embraced the lazy dog." Participants were asked to respond to a multiple choice question on the following page identifying the color of the flamingo. They were unable to return to the previous screen. If they did not select blue, they were rejected from the study. Participants were also given a sample labeling task. They were given emotional statements derived from Havas, Glenberg, and Rinck's emotion comprehension sample set [75]. If participants were unable to identify the emotions in these statements, they were also screened out.

Procedure

A sample of 18 speaker turns were randomly selected from the NYCPHC collection. This collection was chosen over RTFC for this study due to the fact that the entirety of the collection is public and, therefore, a larger sample was eligible to be shared with study participants. In addition, the NYCPHC was similar in scope to RTFC in that it is also a library of community stories. Thus, insights gleaned based on this set could be reasonably applied to RTFC and similar collections.

I filtered the pool of eligible speaker turns to include public turns from the NYCPHC set that were between 25 to 90 seconds in duration. To ensure testers were exposed to turns of similar length.

For this pilot study I did not evaluate all 53 emotions returned by Hume, nor did I evaluate all 16 of my filtered emotion set. I evaluated only a core set of basic emotions based on the Discrete Emotion Questionnaire (DEQ) [76]: Anger, Fear, Happiness, Sadness, and Satisfaction. For my purposes, I determined that if the test were to pass for these core emotions under which many of the emotions in my filtered set fall, I could extrapolate accuracy to the secondary emotions in the filtered set. In the future, a more robust study may be required with more participants and emotions to definitively prove accuracy of AI emotion labeling of the Fora datasets, but this task is slightly out of scope for this thesis. Furthermore, Hume has published several papers and articles validating their methods and can be generally trusted as reliable [62, 77].

For each emotion, I randomly selected a speaker turn of the parameters mentioned that received a score within the following bins: [0.0, 0.25), [0.25, 0.50), [0.25, 0.75), [0.75, 1]. Tables 5-1 and 5-2 represent the full set of speaker turns sampled for this study. Each participant was randomly assigned to an emotion group and presented with the transcripts of 3-4 speaker turns (there were no speaker turns in this collection that scored between 0.75 and 1 for both the emotions Anger and Fear). There were 5 participants per emotion group. They were then given the following prompt for each turn: "Please read the following statement and SELECT ALL of the emotions below that you feel are expressed in the statement." Participants were asked to evaluate the text/language of the speaker turn only due to the limitations of prosody mentioned in the "Underlying Data and Systematic Curation" chapter. Additionally, this was determined to be a comparable task to Hume's language analysis. The emotion options they were presented with were also derived from the DEQ: Anger, Disgust, Fear, Anxiety, Sadness, Desire, Relaxation, and Happiness (relabeled "Joy" in the survey to align with Hume's labels). I decided to use this smaller list instead of the full 53 to avoid overwhelming the participants and confusing the results, especially given the small number of participants. The order of the speaker turns and the order of the emotion choices were

Emotion	Score	Duration (s)	Transcript Excerpt
Anger	0.038	56.333	"...Why is it that some of the elderly are so freaking hardheaded?..."
Anger	0.330	47.635	"Yeah, nebulizer, it did me in. I woke up out of it and I walked around a little bit in the house and got myself together, but in my dream, it finished me..."
Anger	0.691	36.555	"...Now, I have to pay unemployment back $10,000. Misconduct. Do you know what the misconduct was? I didn't take the shot when they wanted me to take the shot, so that turned me very sour against this city. Very."
Fear	0.009	26.07	"Okay, so I'm going to put it on the calendar. Oh, wait, that's the one date..."
Fear	0.260	25.438	"Mine, usually, is 60/110, and my heartbeat is usually 60, 62. No more. I felt there was something wrong with me, and for three hours it was not going anywhere. I was debating, I was so scared. Should I first stay home and die there, or should I go to the hospital and seek help?"
Fear	0.510	65.637	"...I said I wanted to wait until some people take the vaccine and see what the effects were. I though they were using people as guinea pigs..."
Joy	0.0404	36.925	"The laundromat was closed and I didn't have a washing machine at that time or I had a washer, I didn't have a dryer..."
Joy	0.286	56.733	"...You're allowed to walk around it. It's like a park. It's amazing. It's a very beautiful place. So where I was, Prospect Park was open and people took advantage of it and all."
Joy	0.529	47.713	"...I spent days trying to get her an appointment, finally we got through, got her an appointment, her niece took her to that appointment and they were just both crying and so thankful that they could get in..."
Joy	0.803	36.195	"...My name is Leotis. I'm with RiseBoro and we're based in Brooklyn. Today, emotionally, I'm feeling happy, I'm feeling emotionally happy."

Table 5.1: Speaker Turns selected in the Hume Validation Study. Transcripts are truncated above and only display a portion of what study participants saw.

Emotion	Score	Duration (s)	Transcript Excerpt
Sadness	0.0872	73.663	"It's like yo, the masks were a currency. Like, "I'll give you a few cans of bean for a pack of mask." Like that type of situation. You know what I mean? So we're comfortable now. We've dropped our guard..."
Sadness	0.296	32.863	"...I know a lot of the shelters right now, even in New York City, are currently really overwhelmed and a lot of them are seeking services from the public hospitals or the social workers who are themselves also incredibly overwhelmed. So just wanted to drop that in there."
Sadness	0.503	30.915	"When I went to the hospital, for the emergency room, and there was a nurse who took care of me, and she told me that she worked for couple of months in the emergency room for the COVID, and she told me how like 10 nurses that she knew, it was male and female, they died..."
Sadness	0.850	27.88	"Yeah, across from Faber Park. In my situation, we were working right up into the point where they was closed down. I'm a carpenter. Two of our clients had passed away beforehand. One of our clients who was getting ready to start working on their apartment, the lady passed away. She was a younger woman, I believe she was in her 40s-"
Satisfaction	0.0475	51.025	"Thank you. Okay. We can move on to the next section. The next section will be about vaccine experiences and your decision-making. Now I'd like to invite you to share how you made your decision about the COVID-19 vaccine..."
Satisfaction	0.398	28.518	"...But as far as how I feel today, I feel emotionally drained, but spiritually happy. And that's how I'm carrying this day. It could be the weather, but at the end of the day, that's how I feel."
Satisfaction	0.506	41.605	"...So there were different times on the block and every family, the kids would come out and choose their breakfast and go back in. So it took a half hour for all the kids to come out. So we had a half hour break and didn't have to make breakfast. It was something small, but it definitely did something, and the fact that someone took the time to do that."
Satisfaction	0.828	45.237	"... so just knowing that background made me really secure that all these people, these extremely smart people, are looking at the effectiveness, the safety, against other types of vaccines and just knowing that background..."

Table 5.2: Speaker Turns selected in the Hume Validation Study. *Cont.* Transcripts are truncated above and only display a portion of what study participants saw.

randomly displayed. If a participant selected the emotion of their assigned category, it was determined that they were able to detect that emotion at a score within the threshold in question.

Results

Table 5-3 is the summary view of results. If a majority (≥ 3) of the participants were able to detect (select the checkbox of) the emotion at a given threshold, it was determined that a speaker turn with an emotion score with that threshold would generally be accepted by humans as exhibiting that emotion. The results indicate that, if an emotion has a Hume score above 0.5, the participants generally detected that emotion as well.

Emotion	***0.00 - 0.25***	***0.25 - 0.50***	***0.50 - 0.75***	***0.75 - 1.00***
Anger	***True***	False	***True***	x
Fear	False	***True***	***True***	x
Sadness	False	False	***True***	***True***
Joy	False	False	***True***	***True***
Satisfaction*	False	False	False	***True***

Table 5.3: Summary results from the Hume Validation Study. If ≥ 3 particpants recognized the emotions at that threshold, the corresponding cell is marked "True." Participants generally recognized the emotions detected if the associated score was ≥ 0.5. * For Satisfaction, participants were presented with "Relaxation" label from the Discrete Emotion Questionnaire which in the end was not an equivalent proxy.

Discussion

The results of this study led me to establish a threshold of 0.5 for emotion scores. Thus, an emotion label for a speaker turn was only accepted if its score exceeded 0.5, and, during the curation process, turns that did not exhibit an emotion score of at least 0.5 for the filtered set of emotions were excluded. Note that this was a binary task. An evaluation of intensity was not conducted, and intensity is not seriously engaged with in this thesis. Turns were

filtered for minimum detectable levels not for the most intense emotional expression. This filtering only served to ensure the turns selected did indeed match the emotion in question.

5.0.2 Burst Antecedent Curation Study

The second study conducted in this thesis sought to confirm the efficacy of the burst antecedent curation method. The study aimed to answer the following:

- **RQ 1**: Do speaker turns that exhibit the selected emotions and precede vocal bursts lead to stronger reported feelings of associated empathy than randomly selected turns?
- **RQ 2**: Do these curated turns incite emotional disturbance?
- **RQ 3**: Do these turns incite a greater empathy and emotional disturbance than randomly selected turns?

It stands to reason that exposing listeners to emotional or moving audio clips will have an impact on the listener. However, this study verifies whether speaker turns curated with the proposed method are indeed emotionally moving and could, for example, reliably serve as a first pass for collections without highlights.

Participants

Once again, study participants were recruited through the Prolific platform. This study included 150 participants (55% male, 45% female) who were fluent English speakers residing in the US and UK. 81% of the respondents identified as White, 2% as Black, 8% as Asian, and 6% mixed race. Their ages ranged from 18 to 79 ($\mu = 44.2$, $\sigma = 13.877$ years). Participants were paid $3 for their time. The median time taken was 4 minutes, 18 seconds with an average reward per hour of $41.86.

The same attention check screener from the Hume study was used. If a prospective participant was unable to correctly identify the color of the flamingo in the preceding sentence,

they were disqualified. An additional audio check was included where a sentence ("The cow jumped over the moon.") was read aloud and the participants were asked which animal was referenced in the audio clip.

Measures

The intention of *Translations* as a whole and, as a consequence of its underlying curation, is to encourage deep listening. As discussed in the "Background" chapter, the type of listening I hope to inspire (restorative or "reparative") requires empathy and emotional disruption [12, 36]. Emotional disruption can signal that the listener's worldview is being challenged, while empathy facilitates the connection and engagement that is foundational to listening. Thus, if I can demonstrate the preconditions of reparative listening are met, I can move forward to the next phase, hopeful that these stories within the final interface will encourage the listener to linger and listen.

Thus, the measures chosen for this study were associative empathy as represented by Shen's State Empathy Scale [78] and emotional disruption as demonstrated through the Discrete Emotion Questionnaire (DEQ) [76].

Procedure

For this evaluation, I returned to the RTFC collection. I had intended to use the NYCPHC set as well, but, due to unforeseen errors with the AI sensemaking script, I was unable to thematically label enough speaker turns in this collection to run a study that was thematically segmented. The RTFC set was already coded both by AI and humans which would in theory allow me to select non-highlighted speaker turns by theme. However, as previously stated, the RTFC collection is not a fully public collection. Thus, only highlights approved by the participants for public distribution could be used in this study. This is the pool from which I sampled the speaker turns. However, this means that the speaker turns sampled for the study were already part of a curated set: all highlights are selected first by human sensemakers. While these may not necessarily have been selected for my purpose (in the

case of RTFC, sensemakers were instructed to highlight moments where participants merely answered the prompt and shared a question or experience about Boston), they have at least been filtered for relevance and may skew towards emotionally moving stories.

For this study, I compared highlights curated based on their exhibited emotion and whether they preceded vocal bursts with randomly selected highlights. Given that all highlights are curated, I did not expect much of a difference between the two conditions. A follow up study with a fully coded public collection may be conducted in the future.

Once again, I applied a duration filter on the available pool of public turns to ensure that study participants were exposed to audio clips of approximately the same length and that duration of the clip would not be a confounding variable given that story length and time spent listening could also have an impact on the metrics in question. To this end, I filtered for speaker turns within public highlights that were also between 60 and 180 seconds in duration. I then filtered this larger pool for speaker turns labeled with the themes 'Inequality', 'Institutions', and 'Education'.

Theme	Question
Empathy	**E1:** When listening to the story, I was fully absorbed. **E2:** I can relate to what the speaker was going through in the story. **E3:** I can identify with the situation described in the message. **E4:** I can identify with the people in the story.
Emotion	Anger Disgust Fear Anxiety Sadness Desire Relaxation Happiness

Table 5.4: Outline of survey Likert scale survey questions used in the Burst Antecedent Curation Study.

For the control condition, I randomly selected one turn for each of these themes that matched the above parameters, while for the experiment condition I filtered the pool again for turns that preceded vocal bursts and expressed detectable emotions from my filtered set. This

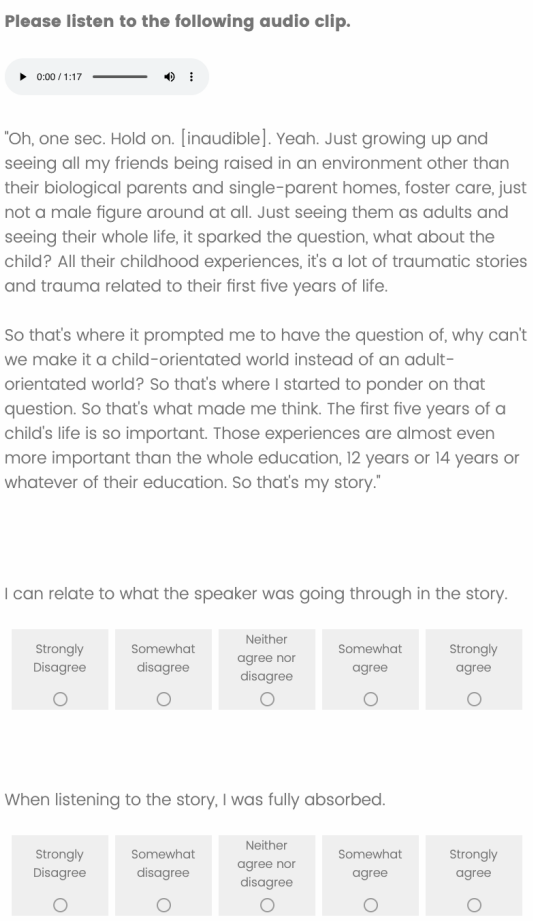

Figure 5-1: Screenshot of speaker turn block presented to participants in the Burst Antecedent Curation Study survey. Participants were presented with the turn audio, transcript, and 4 State Empathy Likert scale questions.

resulted in six conditions: a randomly selected Inequality turn, a randomly selected Institutions turn, a randomly selected Education turn, a curated Inequality turn, a curated Institutions turn, and a curated Education turn. 25 participants were randomly assigned to each condition. Following the screening questions, they were asked to first report their current emotional state on likert scales based on the DEQ. They were then prompted to listen to a single speaker turn based on their assigned condition. The transcript was provided alongside the audio. Below the audio file and the transcript, the four associative empathy likert scale questions were presented.

Condition	Duration (s)	Transcript Excerpt
Random	141.7	"I'll go next. What I wanted to say the question that resonated with me is ironically yours, Fran, because I grew up in Mission Hill during the busing era. I moved here after living in Grove Hall and I came to shell shock because over there I had Black neighbors like myself and then I came here and it was primarily white..."
Random	85.1	"I'll jump and echo, Pastor Reyes sentiment. Everyone has said the same thing, under the larger umbrella of gentrification, and how do we stop that, as... And address particularly Black, unapologetically Black..."
Random	77.2	"Oh, one sec. Hold on. [inaudible]. Yeah. Just growing up and seeing all my friends being raised in an environment other than their biological parents and single-parent homes, foster care, just not a male figure around at all. Just seeing them as adults and seeing their whole life, it sparked the question, what about the child?..."
Curated	81.6	"My question is, how are most people going to ever afford to live in this city? My personal experience with that is that I'm a pastor. I live in a [inaudible], it's not my house. I owned a house in Wisconsin and sold it for a ridiculous price in comparison to here. East Boston was always cheap, and poor, and even though I didn't have any money I could have easily bought a house anywhere around me..."
Curated	113.4	"All right. I have a billion questions. So that was hard to come up with one question. So my question is in regards to marginalized communities and populations, what will their platform look like collectively, more so around expression, real inclusion in policy-making and education?..."
Curated	73.6	"I want to say that you 100% correct in that nothing much has been done to address the issue of mental health with young people. I will take it back to me losing my son. I do have another 16 year old son and I could honestly say nothing was done..."

Table 5.5: Randomly selected speaker turns used in the Burst Antecedent Curation Study. Transcripts are truncated above and only display a portion of what study participants saw.

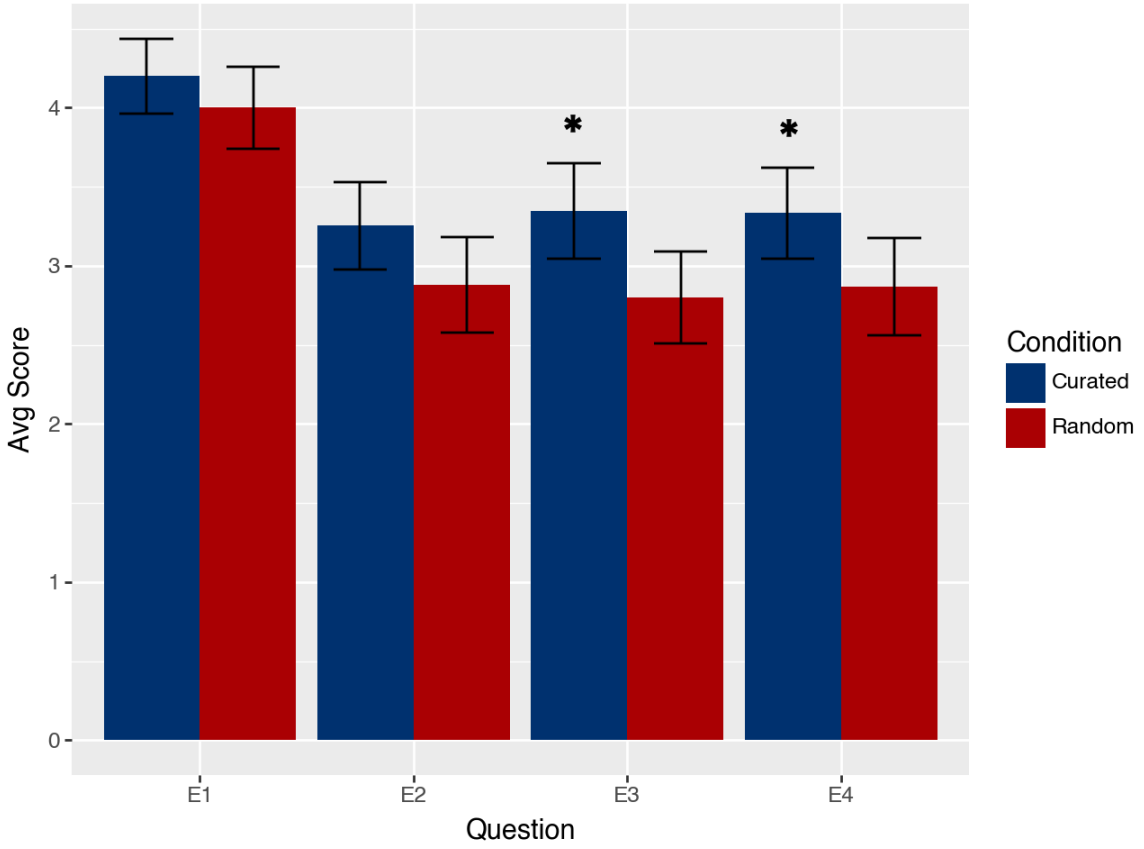

Figure 5-2: Average scores for the 4 State Empathy Likert scale questions for the Burst Antecedent Curation Study. Error bars represent 95% confidence intervals, and * on top of the Curated condition bars indicate $p < 0.05$ for two-sided t-tests comparing against average empathy scores of the Random condition.

Participants were unable to proceed until the audio finished playing and they had provided a response to the State Empathy questions. After listening to the clip and answering the empathy questions, they were asked to record their current emotional state using the DEQ.

Results

Empathy

I conducted two-sided t-tests on the aggregated random selection conditions against the aggregated burst theory curation conditions. I also conducted the same analysis within each theme as well to see if the themes had any bearing on the results. Though the mean likert

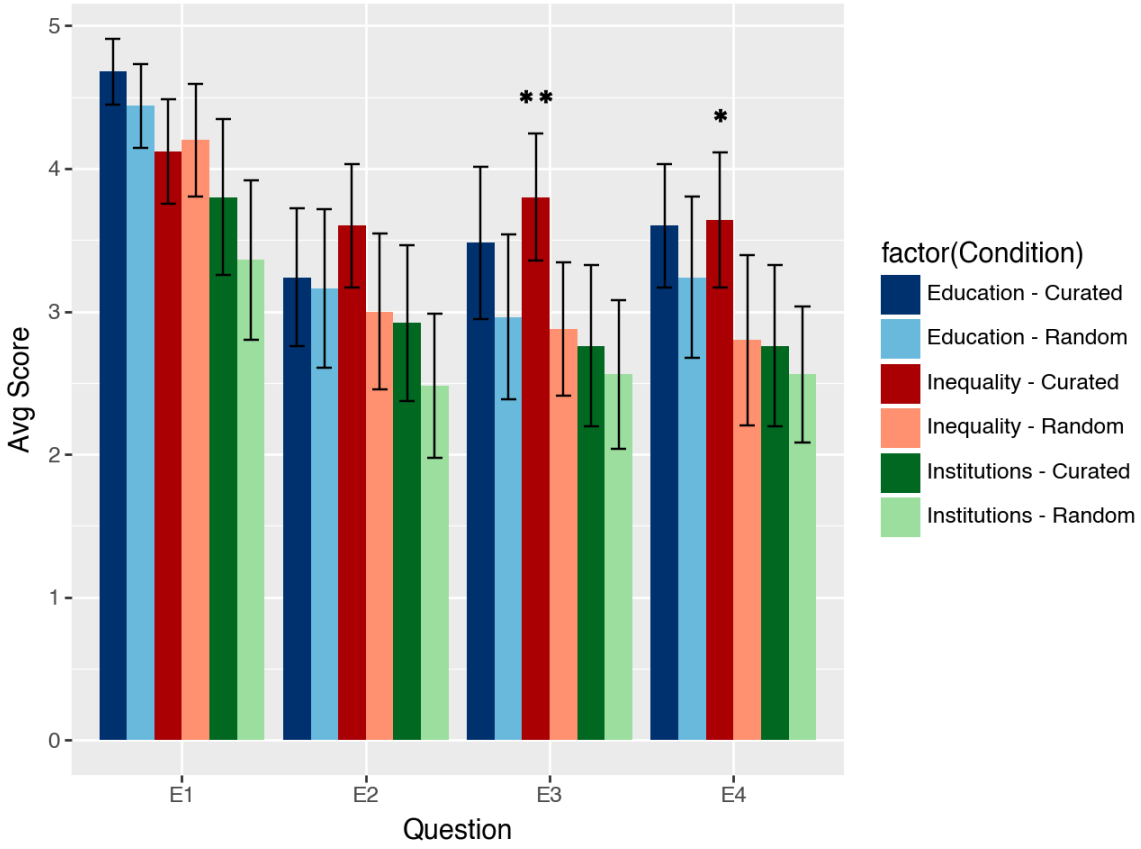

Figure 5-3: Average scores for the 4 State Empathy Likert scale questions for the Burst Antecedent Curation Study, broken out by selected themes: Education, Inequality, and Institutions. Error bars represent 95% confidence intervals, and * and ** on top of the Curated condition bars indicate $p < 0.05$, $p < 0.01$, respectively for two-sided t-tests comparing against average empathy scores of the Random condition.

scale scores of the curated conditions exceeded those of the random condition for each of the associative empathy questions, there was only a significant difference ($p < 0.05$) for statements **E3** and **E4**: "I can identify with the situation described in the message." and "I can identify with the people in the story."

When broken down by theme, it appears that in the case of the theme inequality, there were significant differences ($p < 0.01$) for E3 and ($p < 0.05$) for E4. This suggests that while turns curated with my proposed approach are not significantly better than the random selection of highlights in general, it does appear that this curation method may have merit when applied to certain themes such as those, I suspect, that have a higher volume of speaker

turns. Additional studies would need to be conducted to validate the veracity of this claim.

Emotional Disruption

Figure 5-4 contains slope charts depicting the shift in reported emotions experienced by the participants after exposure to the audio clips. These graphs show that the curated clips did cause a shift in emotion, particularly, an increase in negative emotions and a decrease in positive emotions. Thus, emotional disruption did occur. However, the slopes for each emotion in the curated condition are near parallel to the slopes of the random condition. Therefore, it does not appear that the curated set causes more of a disruption than the random set.

Discussion

The results of this study demonstrate that turns curated with the proposed method do meet the minimum requirements to elicit deep listening. Participants report high associative empathy and experience emotional disruption. However, these curated turns are not consistently, significantly better at eliciting these effects than randomly selected highlights at producing these results. It is important to note, again, that the pool used for evaluation here was a pre-curated set. And, while these systematically curated turns may not beat human curation, it is still possible that they are better than randomly selecting public turns. This method may not be a substitute for human sensemaking (that was never the goal), but it is comparable. Just as it is not significantly better, it is not significantly worse. This suggests that this method could be applied as a first pass for collections or conversations that have yet to undergo or may never undergo human sensemaking or as a supplement to human sensemaking to surface stories that human sensemakers may have missed or excluded.

Due to the time constraints of this thesis, separate evaluations of the effects of turns that express the filtered emotions and those of *any* turns that precede bursts were not conducted, so I cannot determine whether either of these practices or the combination of the two results in the effects described above.

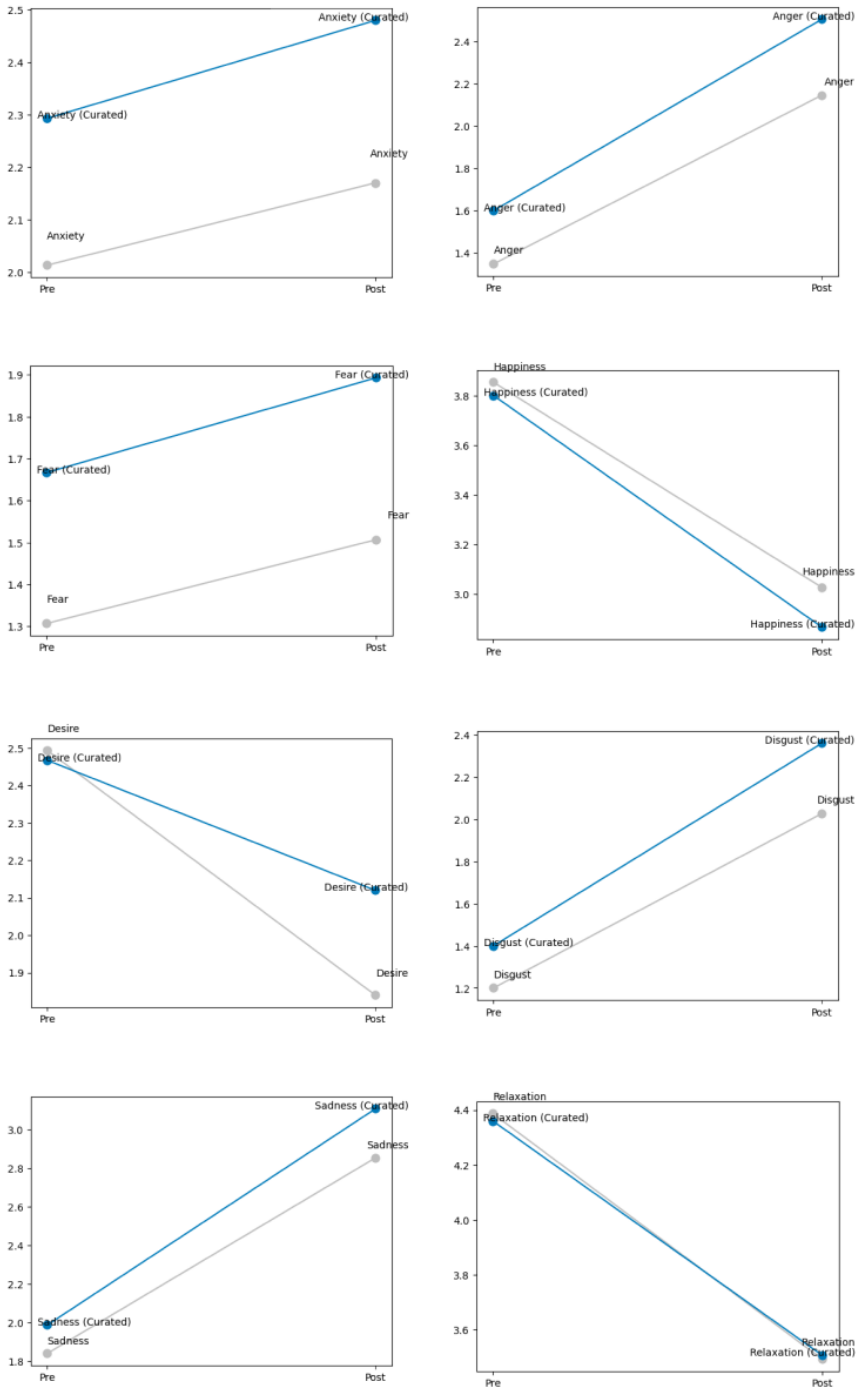

Figure 5-4: Slope charts for the 8 Emotion Likert scale questions from the Discrete Emotions Questionnaire. Participants reported their emotion state before and after exposure to a speaker turn. Near identical levels of emotional disruption occurred across both the Curated and Random conditions, except in the case of Desire. Participants in the Random condition saw a greater decrease in Desire.

5.0.3 The Final Interface

In the final phase of the evaluation, I shared the interface with experts in the media, arts, and anthropology disciplines as well as with friends and family to solicit feedback on the experience as a whole. I intentionally gave them little context to see how they would interact with and interpret the interface without prompting. While there was initial confusion about the spatial organization of the curated speaker turns, those included in this alpha test expressed appreciation of the stories included in the interface. The primary themes of the feedback were layout and organizational improvements, reflections on the AI emotion analysis, reactions to the stories heard, and potential applications of the interface. For each of these user tests, users either interacted with the interface on the public installation or were given a link to the interface and were asked to navigate it with little context. After testing the interface, they shared written feedback, and I conducted over the phone interviews with them as well.

UI Improvements

A common question asked was why the ink drops were laid out in the way they were. One tester asked if the overlap of the drops had significance. In this iteration, any overlap between the drops is coincidental due to the spacing between each conversation line (derived from the number of conversations and the height of the browser window) and the minimum radius I set for the drops (5 pixels). I had preferred the look of the overlap because it mimicked the effect of a splatter of ink drops on printed paper; however, they do suggest clustering. One implementation I had previously considered was clustering the drops by sub themes and shared experiences when the thematic and emotional filters were selected. The feedback provided by this tester indicates that this effect may be a natural affordance. In lieu of such clustering, I could also decrease the base radius of the drops, thereby implying relation only to their respective conversation line instead of with the drops from other conversations. Relationship and connection with other stories in the collection could then be revealed on hover or when the user selected a theme.

In that vein, another tester, a UI designer suggested that the drops appear in reaction to the user's mouse movements and other interactions, forcing the user to more actively search for stories. This effect would be akin to the physical motion required in "Stories in the Rain." By nudging the user to make active choices, their buy in and attention to the selected stories would increase.

Furthermore, a criticism received was that 82 highlights displayed were still too many, and that the user's view could still be more limited. One method of doing so would be to start with an even more limited set of curated turns to start and entice the user and then slowly reveal more turns as they listened.

Testers often asked the meaning behind the xy-layout of the drops. Some inquired if the stories were ranked based on the intensity of the emotion, for example, or if the horizontal position was related to the themes listed across the top. As previously explained, the thematic and emotional filters currently have no bearing on the spatial layout. The vertical axis reflects the order in which the conversations were held, while the horizontal axis reflects the proportional time within the conversations. These tester reactions suggest that more spacing can be provided between the filters and the graph to avoid users misconstruing them as axis labels, and separate axis labels could be included. In fact, the design mockups for the final interface did include such labels but the alpha tested version did not. The user feedback emphasized the importance of such labels. Many users also failed to recognize that each line was a conversation. I intended to rely on the animation to communicate that but the alpha version with limited and abbreviated animations did not accomplish that. This feedback has caused me to consider additional annotations to supplement the animations even in the cases where the full animations are present. More radically, the intuition of several of the testers was that the drops were not organized by conversation. They assumed cross-collection associations. This suggests the possibility that the conversation lines could be abandoned entirely. I had included these lines for context–to communicate the source of the drops–but perhaps this context is not necessary throughout the experience.

I observed an interesting phenomenon from two of the testers. They were initially unsure how to engage with the interface after clicking start session. The start page includes an

auto preview of the stories in the collection. One speaker turn is dropped at a time and the user is able to hear 5 seconds of its audio before the next speaker turn starts. After starting the session, the users noted that the audio stopped playing and were unsure how to proceed. They were unaware that they had to click to interact with the second page and continue listening. Other testers did not experience the same confusion, and, when asked about it, they explained that the hover reactions (the opacity and color of the page's elements changing) were enough to signal that the page was interactive. These testers did note, however, that they did not spend much time on the start page, so the expectation of autoplay was not established. The confusion of the two testers did make me reconsider incorporating an autoplay feature in the main interaction as well for those who are less inclined to explore on their own.

Emotion Analysis

Three of the testers called out the accuracy of the emotional analysis with one user stating:

> "I think the AI take on emotion is a little wonky, from what I saw. I clicked through all of the "Infrastructure" stories and the saddest one was color coded as disappointment."

Another tester called out the same speaker turn; however, they noted that (1) though it was not labeled as they would have, they understood Hume's reasoning, and (2) the seeming inaccuracy prompted self-reflection. Knowing the AI provided a label as she listened, she more carefully considered the underlying sentiments and intention of the speaker turns more carefully. This frame of mind allowed her to better empathize with the speaker and reflect on her own understanding and connection with the experience.

Another user noted that, while searching for joy in the collection, they came across a speaker turn that as a whole was not joyful (the speaker was sharing their experience with gun violence), but the turn had been labeled with joy because the speaker ended the turn saying she is happy her husband survived. Such contradictions are common with language-based

analysis, and this is an area where the interface would benefit from other supplemental forms of emotional analysis. For the goal of curating moving stories, perhaps these "mistakes" are ok, but one of my intentions is that the user have control over the types of stories they wish to hear. If they are seeking joy, they should find joy and not be surprised with traumatic stories.

Curation

All testers expressed that they had been moved by the stories they heard. One shared:

> "some of these are really powerful. . . just listened to a woman talk about how it's impossible for minorities to own in boston, and even making rent requires 2 minimum wage jobs. damn."

Applications

The testers encompassed professionals from several industries including journalism, media consulting, Social Media, and music. One tester reflected that such an interface would be useful in the coming US presidential election cycle and could capture reactions to presidential debates, while others shared:

> "It feels a bit like something you might find in a museum – here's a bunch of people talking about their experiences and this might help you decide what you want to listen to first, given that you probably don't have the time to listen to all of them (and you may not be as interested in some of the topics)."
>
> "I could see this as a way for a documentarian to organize a bunch of clips. I could also see mood as a way of organizing songs – I know Spotify uses mood to organize some of its playlists. Mood is also something used by good DJs to figure out their set of music and the flow between songs – usually to ramp up or down to the extremes rather than playing the extremes right next to each other. DJs use tempo and familiar/unfamiliar, too, it's not primarily mood."

Chapter 6

Conclusion

6.0.1 Takeaways and Future Work

There are several unexplored avenues for future work. In fact, during the process of developing this thesis, it was difficult to resist the urge to follow the many threads that appeared.

Installation

I had set out to design an engaging interaction, but, while the final interface incorporated thoughtful and intentional motion design and some reveal effects in response to user clicks, I did not have the opportunity to fully realize this vision. The lowest hanging fruit for revision in this area is to incorporate the motion sense used in "Stories in the Rain." With an interface where the curation, duration of play, and selection method is sorted out, the physical motion from "Stories" would enrich this experience.

Another point of improvement is further limiting the field of view.The current interface narrows approximately 11K speaker turns and 1300 highlights to 82 curated turns. However, this is still too many. Stricter thresholds could be applied. In the end, given the limitations I had with displaying the full conversation audio, I opted for more highlights and lowered the burst threshold to 0.3. Increasing it to the suggested threshold of 0.5 would further limit

the set. Starting from a limited set and encouraging playful exploration by slowly revealing related, available speaker turns would also keep users' attention longer.

In the end, on the spectrum between form and function, Translations sits closer to the side of function than I initially intended. I leaned more heavily towards a visualization of emotions in the collection than towards a cohesive audio-visual experience. Beyond the interaction design, I was unable to fully realize my vision for the audio design of the piece. Playing the burst reactions to turns you hover over first, for example, to tease interesting moments or aligning similar words and phrases within the collection to emphasize consensus through the ear. This thesis sits in the middle. It is neither a standalone public installation nor a research or analysis tool. It is a playful explorer that exposes stories to the public, but, without the constraints of the thesis, I am excited to explore both worlds: fully artistic translations of voice and participant dynamics as well as functional tools that reveal insights about participant behavior that may allow us to improve the design of communication spaces.

Functional Analysis

As I have stated, throughout this process and much of my previous work, there has been a tension between form and function. The initial proposal for this thesis was a more playful interaction for the public and that is what I achieved. But, while collecting data for this interface, I stumbled upon many interesting questions. For example, what are the ways in which individuals dominate conversations (beyond speaking time) and how do we mitigate these? How are our facilitators affecting the dynamics of the conversation? These questions are becoming easier to answer with tools such as Hume and others like the Linguistic Inquiry Word Count (LIWC), a tool I tested but was unable to deploy in this incarnation of the thesis [79, 80].

There are choices that I could have made to center these analyses and questions. For example, adding a toggle to rank conversations based on facilitator intervention. I opted not to do so in favor of public interest over research motivations. But, these are still valid modes of inquiry I hope to explore in future work. Ultimately, my goal is to apply these tools

and analyses in the media space. I wish to illustrate how various media and communication spaces affect us. Such tools would be helpful in that quest. Additionally, within the CCC and Fora context, the analyses I conducted throughout this thesis and the tools that could be built from them could be used to evaluate if we are living up to our mission. In the final interface, there were no joyful highlights. Are these the type of stories we wish to surface? Perhaps in the case of RTFC these are sufficient since the purpose was to surface pain points and action items. But, maybe we do wish to include productive joy in our interactions.

Zoom and Virtual Interaction

A non-trivial portion of the conversations in both of the collections referenced in this thesis were held over Zoom. Meeting in the virtual space has an impact on participant behavior. For example, we often mute ourselves or are less engaged. These actions would have an impact on the available audio cues such as vocal bursts. For this thesis, I did not account for these differences, yet, in the RTFC set at least there appear to be enough bursts for curation. However, more careful consideration of what participant signals can be used for virtual conversations is necessary.

Curation

Translations utilized fixed thresholds for the emotion analysis, but dynamic thresholds attuned to the behaviors in a collection for each emotion could be applied. In the Hume validation pilot study, 0.5 was identified as a reliable threshold across emotions, but Fear was detected within the [0.25-0.50) range and Anger was detected at [0.0, 0.25). A more responsive system would allow for a richer curation.

Emotional Complexity and Expression

Human expression of emotion is complex. I agree with Hume's premise that we should acknowledge and express the vastness of human expression, allowing observers to make

nuanced inferences. I did not wish to condense this complexity into flat states. For simplicity and time, I selected the top emotions for visualization, but I still believe a design that reflects the fluidity and complexity of human emotion is possible.

6.0.2 Reflections

> "The longer I live, the more deeply I learn that love — whether we call it friendship or family or romance — is the work of mirroring and magnifying each other's light. Gentle work. Steadfast work. Life-saving work in those moments when life and shame and sorrow occlude our own light from our view, but there is still a clear-eyed loving person to beam it back. In our best moments, we are that person for another."
>
> *James Baldwin*

There were three inciting incidents that precipitated this thesis. The first happened back in 2016. Amidst national discussion about police brutality, ABC's Black-ish aired an episode titled "Hope." This was a bottle episode during which the Johnson family, gathered around the television, watches a news report on the over-policing of and institutional violence against black bodies and the ensuing protests. The entire episode is a conversation among a black family. At the time, I did not deem this as radical, but news outlets such as Vox and Slate did [81, 82]. They lauded the episode not only for talking candidly about the sensitive topics but also for modeling these conversations for audiences who had never experienced such discussions or were even aware that these conversations were happening in most if not all black homes across the US. Through dialogue–even scripted–the Johnson family had an impact on those for whom such perspectives were previously inaccessible, or so these outlets claimed. I was–and still am–curious if that was true. Did this fictional dialogue actually

have a positive–or any–impact on public discourse? How could we determine this beyond anecdotal evidence? This is a question I hope someday to answer.

The second inciting incident occurred more recently (2023): the alleged downfall of Black Twitter, which has been a source of culture for over a decade. It is a place on the internet that has brought me and many others joy and commiseration. In recent years, prior even to Musk's acquisition, it has become a less fun place. I recall, following the backlash to the wrongful death of George Floyd, a viral video of a black woman singing a parody of "What a friend we have in Jesus" surfaced. Instead of the titular lyrics, she sang "What a friend we have in Google. All our questions answered there. What a privilege to Google. Instead of asking our black friends." (To my surprise, upon attempting to relocate this video for citation, I discovered this new spiritual was also featured an episode from a later season of Black-ish, Season 7, Episode 10: "What About Gary?"). In this song, the frustrations of being expected to educate and validate our peers is captured. The words acknowledge the emotional labor involved in such pursuits. During this time, many experienced a form of dominance not often talked about. In the "Introduction" chapter I mention hate speech, but even when we are not actively being bullied or explicitly harmed and are perhaps seemingly given a space to share, we are not given space to breathe. The expectations placed on us to share traumatic stories is suffocating. I now know that this is ethical loneliness and that we lack reparative and restorative listening. I did not know this then and was met with a sense of loss. I have carried this sensation with me since then, in fact, into the thesis process. I had hoped to perform a data analysis on Twitter or other media platforms to quantify this feeling. Though the final thesis proposal was an interface, I at least wanted to facilitate restorative listening and commiseration, and I sought metrics that could potentially flag the problematic phenomena that I had experienced since 2021.

The third inciting incident was when I listened to Real Talk for Change Boston conversation 36 for the *Conversation Prints* project. I was well into my first semester but was listening to a conversation in full for the first time. While logging sounds and the absence of sounds, I heard powerful stories. What's more, when I visited other conversations in the collection I found some had no highlights. I was filled with disappointment that many would never hear

the direct accounts in these collections. Ever since, I have taken a bit of a detour from my intended career trajectory of media analytics and system design to explore how to create public experiences that facilitate engagement with conversation corpuses.

From *Conversation Prints*, I learned the utility of natural metaphors. I became adept at p5.js and was reminded of the importance of audio in moving an audience. Through "Theme Sort," I began to fully explore the richness of conversation data and navigated the challenges of representing collection level data. *Collective Echos* showed me the power of physical space and that movement could be a proxy for active listening and engagement. *Collective Echos* also highlighted the power of curation. Finally, “Stories in the Rain” allowed me to further test alternate modes of interaction and the effects of motion graphics in this context.

These learnings and others culminated in *Translations*. This project has laid the foundation necessary for future development of the ideas established here. I may not yet know exactly how Black-ish affected public discourse, but I am now able to detect emotional response within conversations at a level of granularity that would allow for impact measurement. And, while there was not enough demographic data for comprehensive analysis across an entire corpus, I can detect and visualize to some degree the emotional experience of people of color within a conversation collection.

When I first arrived at the Media Lab, I mentioned to my advisor that I was intrigued by the idea of automatic audio medley curation. At the time, I never envisioned that I would develop a framework for curation. These were not skills I possessed, and I had only hoped to design the wrappers around such tools. In pursuit of *Translations*, I have offered a framework for systematic curation based on participant reactions that elicits the necessary empathetic and emotional response for restorative listening. Simultaneously, I present an interface that allows listeners to explore these curated stories. *Translations* provides a playful and insightful means of exploration through our corpus. I am extremely honored to have had the opportunity to care for and share the stories in the RTFC collection these last two years. And, I am excited to continue to explore conversations as an artistic and evaluative instrument.

Bibliography

[1] E. Dwoskin, "Opinion | Why Black Twitter will never die," Jul. 2023. [Online]. Available: https://www.msnbc.com/opinion/msnbc-opinion/elon-musk-black-twitter-future-rcna93458

[2] W. Hood and G. M. Tada, Eds., *Black Landscapes Matter*. University of Virginia Press, 2020.

[3] M. D. Clark, Ph.D., "Archiving Black Twitter." [Online]. Available: https://www.meredithdclark.com/archivingblacktwitter

[4] "SOMEWHERE GOOD." [Online]. Available: https://www.somewheregood.com/

[5] "Spill." [Online]. Available: https://www.spill.com/

[6] D. Roy, "The Internet Could Be So Good. Really." Oct. 2023, section: Technology. [Online]. Available: https://www.theatlantic.com/technology/archive/2023/10/social-media-platforms-business-models-dialogue-network/675655/

[7] "Real Talk for Change." [Online]. Available: https://realtalkforchange.org/

[8] M. Hughes, "K E E P E R : online conversation support scaffolding modeled after ancient and modern social technologies," Thesis, Massachusetts Institute of Technology, 2020, accepted: 2021-05-14T16:28:51Z Journal Abbreviation: Online conversation support scaffolding modeled after ancient and modern social technologies. [Online]. Available: https://dspace.mit.edu/handle/1721.1/130606

[9] A. Errante, "But Sometimes You're Not Part of the Story: Oral Histories and Ways of Remembering and Telling," *Educational Researcher*, vol. 29, no. 2, pp. 16–27, Mar. 2000, publisher: American Educational Research Association. [Online]. Available: https://doi.org/10.3102/0013189X029002016

[10] C. A. Riley, *This Here Flesh*. Penguin Random House, Jan. 2023.

[11] "Fora • Conversation Platform." [Online]. Available: https://fora.io/

[12] S. F. Harding, "Convicted by the Holy Spirit: The Rhetoric of Fundamental Baptist Conversion," *American Ethnologist*, vol. 14, no. 1, pp. 167–181, 1987, publisher: [Wiley, American Anthropological Association]. [Online]. Available: https://www.jstor.org/stable/645639

[13] "Group Overview ‹ Social Machines." [Online]. Available: https://www.media.mit.edu/groups/social-machines/overview/

[14] N. Gillani, A. Yuan, M. Saveski, S. Vosoughi, and D. Roy, "Me, My Echo Chamber, and I: Introspection on Social Media Polarization," in *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, 2018, pp. 823–831, arXiv:1803.01731 [cs]. [Online]. Available: http://arxiv.org/abs/1803.01731

[15] M. Saveski, "Polarization and toxicity in political discourse online," Thesis, Massachusetts Institute of Technology, 2020, accepted: 2021-01-06T20:18:36Z. [Online]. Available: https://dspace.mit.edu/handle/1721.1/129320

[16] P. Beshai, "Visualizing Toxicity in Twitter Conversations," Aug. 2018. [Online]. Available: https://medium.com/cortico/visualizing-toxicity-in-twitter-conversations-3cd336e5db81

[17] C. A. Bail, L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky, "Exposure to opposing views on social media can increase political polarization," *Proceedings of the National Academy of Sciences*, vol. 115, no. 37, pp. 9216–9221, Sep. 2018,

publisher: Proceedings of the National Academy of Sciences. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.1804840115

[18] C. Bail, "Breaking the Social Media Prism: How to Make Our Platforms Less Polarizing," in *Breaking the Social Media Prism*. Princeton University Press, Sep. 2022. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/9780691246499/html

[19] M. Saveski, B. Roy, and D. Roy, "The Structure of Toxic Conversations on Twitter," in *Proceedings of the Web Conference 2021*, ser. WWW '21. New York, NY, USA: Association for Computing Machinery, Jun. 2021, pp. 1086–1097. [Online]. Available: https://dl.acm.org/doi/10.1145/3442381.3449861

[20] Cortico, "Local voices network - about." [Online]. Available: https://lvn.org/about/

[21] ——, "_USA New Era of In-Person Conversations - Cortico," Oct. 2023. [Online]. Available: https://cortico.ai/news/product-updates/a-new-era-of-in-person-conversations/

[22] ——, "_USSensemaking at Cortico - Cortico," Nov. 2023. [Online]. Available: https://cortico.ai/news/sensemaking-at-cortico/

[23] sscott, "Making sense of sensemaking: What it is and what it means for pandemic research," Aug. 2020. [Online]. Available: https://www.atlanticcouncil.org/blogs/geotech-cues/making-sense-of-sensemaking-what-it-is-and-what-it-means-for-pandemic-research/

[24] "Real Talk for Change Sensemaking Report," MIT Center for Constructive Communication, Tech. Rep., 2021.

[25] "GLOBAL WE," Jul. 2023. [Online]. Available: https://www.museumfortheunitednations.com/global-we

[26] Cortico, "_USA New Era of In-Person Conversations - Cortico," Oct. 2023. [Online]. Available: https://cortico.ai/news/product-updates/a-new-era-of-in-person-conversations/

[27] ——, "_USIntroducing Fora - Cortico," Sep. 2023. [Online]. Available: https://cortico.ai/news/fora/

[28] "Youth Voices — JA Worldwide." [Online]. Available: https://www.youthvoices.org/

[29] M. A. Hughes and D. Roy, "Keeper: A Synchronous Online Conversation Environment Informed by In-Person Facilitation Practices," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, May 2021, pp. 1–14. [Online]. Available: https://dl.acm.org/doi/10.1145/3411764.3445316

[30] C. Overney, "SenseMate: An AI-Based Platform to Support Qualitative Coding," Thesis, Massachusetts Institute of Technology, Jun. 2023, accepted: 2023-08-30T15:56:08Z. [Online]. Available: https://dspace.mit.edu/handle/1721.1/151980

[31] M. Hughes, A. Luk, N. Obeng-Marnu, and D. Dimitrakopoulou, "Interactive Visualizations of Personal Experiences for Shared Sense of Understanding," *Commonplace*, vol. 3, no. 1, Jun. 2023. [Online]. Available: https://commonplace.knowledgefutures.org/pub/rvtt9ex4/release/1

[32] M. Hughes, B. Roy, and D. Roy, "Project Overview ‹ Conversation Dynamics." [Online]. Available: https://www.media.mit.edu/projects/conversation-dynamics/overview/

[33] ——, "Project Overview ‹ Conversation Chains." [Online]. Available: https://www.media.mit.edu/projects/conversation-chains/overview/

[34] L. Goubert, K. D. Craig, T. Vervoort, S. Morley, M. J. L. Sullivan, d. C. A. C. Williams, A. Cano, and G. Crombez, "Facing others in pain: the effects of empathy," *PAIN*, vol. 118, no. 3, p. 285, Dec. 2005.

[35] P. Wright and J. McCarthy, "Empathy and experience in HCI," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '08. New York, NY, USA: Association for Computing Machinery, Apr. 2008, pp. 637–646. [Online]. Available: https://dl.acm.org/doi/10.1145/1357054.1357156

[36] J. Stauffer, *Ethical Loneliness: The Injustice of Not Being Heard.* Columbia University Press, Sep. 2015, pages: 240 Pages.

[37] C. Menkel-Meadow, "Restorative Justice: What Is It and Does It Work?" *Annual Review of Law and Social Science*, vol. 3, no. Volume 3, 2007, pp. 161–187, Dec. 2007, publisher: Annual Reviews. [Online]. Available: https://www.annualreviews.org/content/journals/10.1146/annurev.lawsocsci.2.081805.110005

[38] H. R. Bernard, *Research Methods in Anthropology: Qualitative and Quantitative Approaches*, 6th ed. Rowman & Littlefield, 2018, google-Books-ID: JzSLswEACAAJ.

[39] K. Charmaz, *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis.* SAGE, Jan. 2006, google-Books-ID: v1qP1KbXz1AC.

[40] "NVivo." [Online]. Available: https://lumivero.com/products/nvivo/

[41] S. Chandrasegaran, C. Bryan, H. Shidara, T. Chuang, and K. Ma, "TalkTraces | Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems." [Online]. Available: https://dl.acm.org/doi/10.1145/3290605.3300807

[42] J. Donath and d. boyd, "Public Displays of Connection," *Bt Technology Journal - BT TECHNOL J*, vol. 22, pp. 71–82, Oct. 2004.

[43] J. Donath, "Developing Legible Visualizations for Online Social Spaces," 2002. [Online]. Available: https://www.academia.edu/download/51358837/Developing_Legible_Visualizations_for_On20170115-19001-p918hv.pdf

[44] J. Donath, K. Karahalios, and F. Viegas, "Visualizing conversation," in *Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences. 1999. HICSS-32. Abstracts and CD-ROM of Full Papers*, vol. Track2, Jan. 1999, pp. 9 pp.–. [Online]. Available: https://ieeexplore.ieee.org/document/772675

[45] "About — Binaura." [Online]. Available: https://www.binaura.net/

[46] "Bricolage — Binaura." [Online]. Available: https://www.binaura.net/bricolage

[47] H. Hansen, "THE VALUE OF A LINE." [Online]. Available: https://www.studioheatherhansen.com/portfolio-3/the-value-of-a-line

[48] "Guillaume Wiatr | The Art of Visualizing Conversations." [Online]. Available: https://paperlike.com/blogs/paperlikers-insights/guillaume-wiatr-the-art-of-visualizing-conversations-1

[49] "DRAWN TO LEAD." [Online]. Available: https://www.drawntolead.org/

[50] G. Lupi, "Data Humanism, the Revolution will be Visualized." Jul. 2017. [Online]. Available: https://medium.com/@giorgialupi/data-humanism-the-revolution-will-be-visualized-31486a30dbfb

[51] ——, "Unraveled Stories." [Online]. Available: http://giorgialupi.com/unraveled-stories

[52] S. Posavec, "Papworth Hospital, Inpatients Ward," 2019.

[53] P. Palmer, "Laughter & Silence: The Spiritual Odd Couple | The On Being Project." [Online]. Available: https://onbeing.org/blog/laughter-silence-the-spiritual-odd-couple/

[54] T. L. Chartrand and J. A. Bargh, "The chameleon effect: The perception–behavior link and social interaction," *Journal of Personality and Social Psychology*, vol. 76, no. 6, pp. 893–910, 1999, place: US Publisher: American Psychological Association.

[55] E. Roth, "Multi-Touch Paintings Series," 2014.

[56] B. Fry, "On the Origin of Species: The Preservation of Favoured Traces," 2009.

[57] L. Nicoletti and S. Sarva, "When Women Make Headlines." [Online]. Available: https://pudding.cool/2022/02/women-in-headlines/

[58] M. Hughes, C. Lee, N. Obeng-Marnu, and S. Russell, "Collective Echos."

[59] J. Cardiff and G. B. Miller, "Osaka Symphony," 2018.

[60] "Cortico Researcher API." [Online]. Available: https://api.fora.io/docsget-/v1/conversations

[61] "Hume Expression Measurement API." [Online]. Available: https://www.hume.ai/productsexpressionMeasurementAPI

[62] A. S. Cowen, D. Keltner, F. Schroff, B. Jou, H. Adam, and G. Prasad, "Sixteen facial expressions occur in similar contexts worldwide," *Nature*, vol. 589, no. 7841, pp. 251–257, Jan. 2021, publisher: Nature Publishing Group. [Online]. Available: https://www.nature.com/articles/s41586-020-3037-7

[63] A. Cowen, D. Sauter, J. L. Tracy, and D. Keltner, "Mapping the Passions: Toward a High-Dimensional Taxonomy of Emotional Experience and Expression," 2019. [Online]. Available: https://journals.sagepub.com/doi/epub/10.1177/1529100619850176

[64] A. S. Cowen and D. Keltner, "Self-report captures 27 distinct categories of emotion bridged by continuous gradients," *Proceedings of the National Academy of Sciences*, vol. 114, no. 38, pp. E7900–E7909, Sep. 2017, publisher: Proceedings of the National Academy of Sciences. [Online]. Available: https://www.pnas.org/doi/10.1073/pnas.1702247114

[65] "Expression Measurement API – Hume API." [Online]. Available: https://dev.hume.ai/docs/expression-measurement-api/overview

[66] A. Celle, *Expressing and Describing Surprise*. John Benjamins Publishing Company, publication Title: bct.92. [Online]. Available: https://benjamins.com/catalog/bct.92

[67] K. Hengeveld and H. Olbertz, "Didn't you know? Mirativity does exist!" *Linguistic Typology*, vol. 16, no. 3, pp. 487–503, Nov. 2012, publisher: De Gruyter Mouton. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/lity-2012-0018/html

[68] A. Y. Aikhenvald, "The essence of mirativity," *Linguistic Typology*, vol. 16, no. 3, pp. 435–485, Nov. 2012, publisher: De Gruyter Mouton. [Online]. Available: https://www.degruyter.com/document/doi/10.1515/lity-2012-0017/html

[69] T. M. Sutton and J. Altarriba, "Color associations to emotion and emotion-laden words: A collection of norms for stimulus construction and selection," *Behavior*

Research Methods, vol. 48, no. 2, pp. 686–728, Jun. 2016. [Online]. Available: https://doi.org/10.3758/s13428-015-0598-8

[70] W. R. Crozier, "Shame, Guilt and Anger," in *Blushing and the Social Emotions: The Self Unmasked*, W. R. Crozier, Ed. London: Palgrave Macmillan UK, 2006, pp. 134–148. [Online]. Available: https://doi.org/10.1057/9780230501942_8

[71] D. Jacquette, "Hume on Color Knowledge, with an Application to Jackson's Thought Experiment," *History of Philosophy Quarterly*, vol. 29, no. 4, pp. 355–371, 2012, publisher: [North American Philosophical Publications, University of Illinois Press]. [Online]. Available: https://www.jstor.org/stable/43488049

[72] A. N. Gilbert, A. J. Fridlund, and L. A. Lucchina, "The color of emotion: A metric for implicit color associations," *Food Quality and Preference*, vol. 52, pp. 203–210, 2016, place: Netherlands Publisher: Elsevier Science.

[73] L. Zaikauskaite, P. French, M. Stojanovic, and D. Tsivrikos, "The effects of moral context on the colours of guilt and pride," *The Social Science Journal*, vol. 60, no. 4, pp. 802–817, Oct. 2023, publisher: Routledge _eprint: https://doi.org/10.1080/03623319.2020.1782637. [Online]. Available: https://doi.org/10.1080/03623319.2020.1782637

[74] G. Kleinert, ""The Köln Concert part II c" Keith Jarrett," 2018.

[75] D. A. Havas, A. M. Glenberg, and M. Rinck, "Emotion simulation during language comprehension," *Psychonomic Bulletin & Review*, vol. 14, no. 3, pp. 436–441, Jun. 2007. [Online]. Available: https://doi.org/10.3758/BF03194085

[76] C. Harmon-Jones, B. Bastian, and E. Harmon-Jones, "The Discrete Emotions Questionnaire: A New Tool for Measuring State Self-Reported Emotions," *PLoS ONE*, vol. 11, no. 8, p. e0159915, Aug. 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4976910/

[77] J. Brooks, PhD, "What is semantic space theory?" [Online]. Available: https://www.hume.ai/blog/what-is-semantic-space-theory

[78] L. Shen, "On a Scale of State Empathy During Message Processing," *Western Journal of Communication*, vol. 74, no. 5, pp. 504–524, Oct. 2010, publisher: Routledge _eprint: https://doi.org/10.1080/10570314.2010.512278. [Online]. Available: https://doi.org/10.1080/10570314.2010.512278

[79] "Welcome to LIWC-22." [Online]. Available: https://www.liwc.app/

[80] Ryan L. Boyd, Ashwini Ashokkumar, Sarah Seraj, and James W. Pennebaker, "The Development and Psychometric Properties of LIWC-22," The University of Texas at Austin, Tech. Rep., 2022. [Online]. Available: https://www.liwc.app

[81] A. Harris, "Black-ish Just Delivered One of the Best Episodes About Race Ever to Air on Television," *Slate*, Feb. 2016, section: Brow Beat. [Online]. Available: https://slate.com/culture/2016/02/black-ishs-episode-on-police-brutality-hope-is-one-of-televisions-best-depictions-of-police-brutality-yet-video.html

[82] C. Framke, "Black-ish lets the family talk candidly about police brutality - Vox," Feb. 2016. [Online]. Available: https://www.vox.com/2016/2/28/11129642/blackish-police-brutality-episode-review