CS147 - Mixed Reality Studio

Assignment 5
Low-fi Prototypes

Team

Abdallah A.
Computer

Ahmed S.
Facilitator

Ammar A.
Note-taker

Alejandrina G. R.
Note-taker

Introduction

Value Proposition

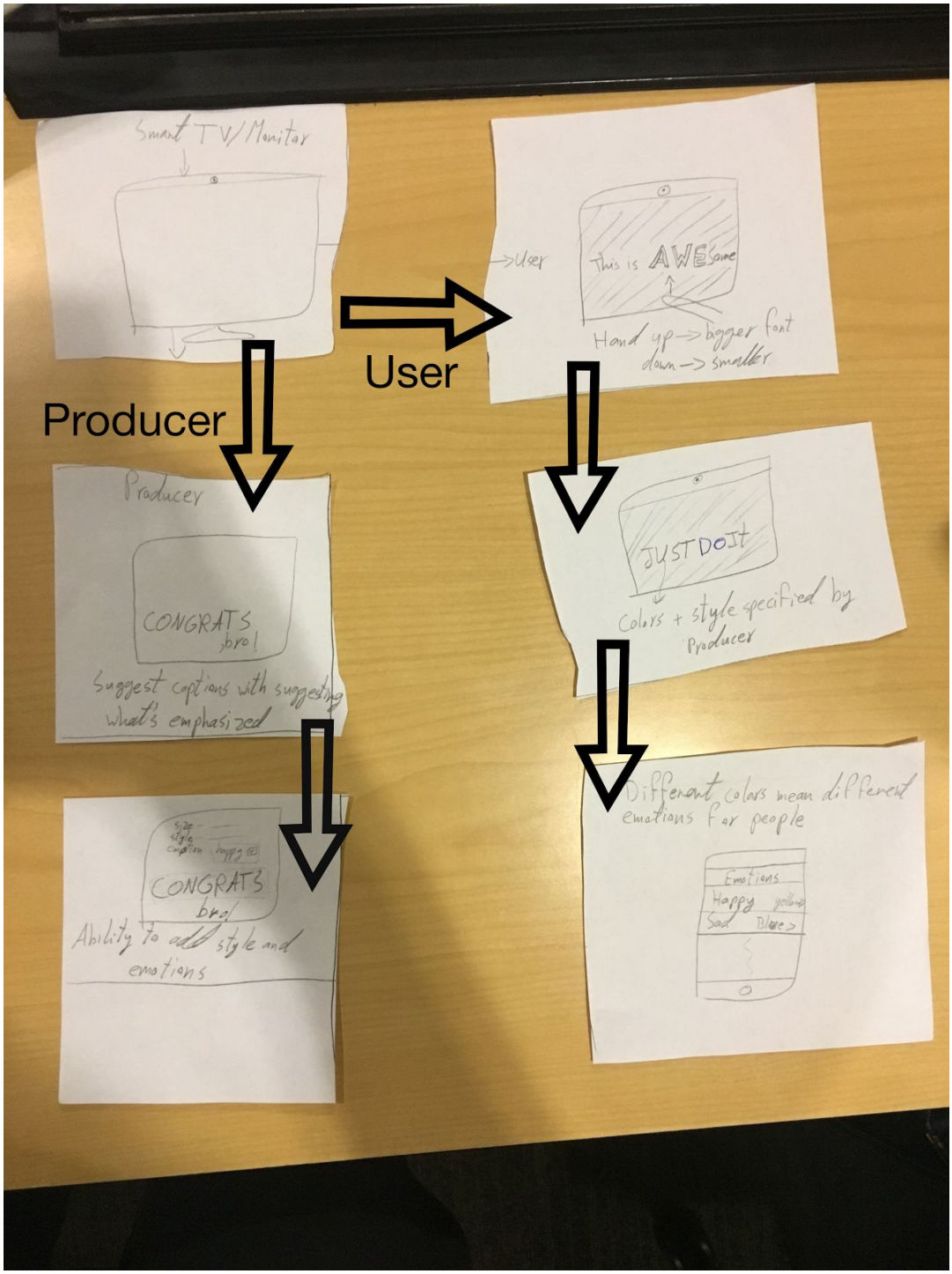
Conveying the tone, volume, and emotion implicit in speech through the medium of text.

Problem & Solution

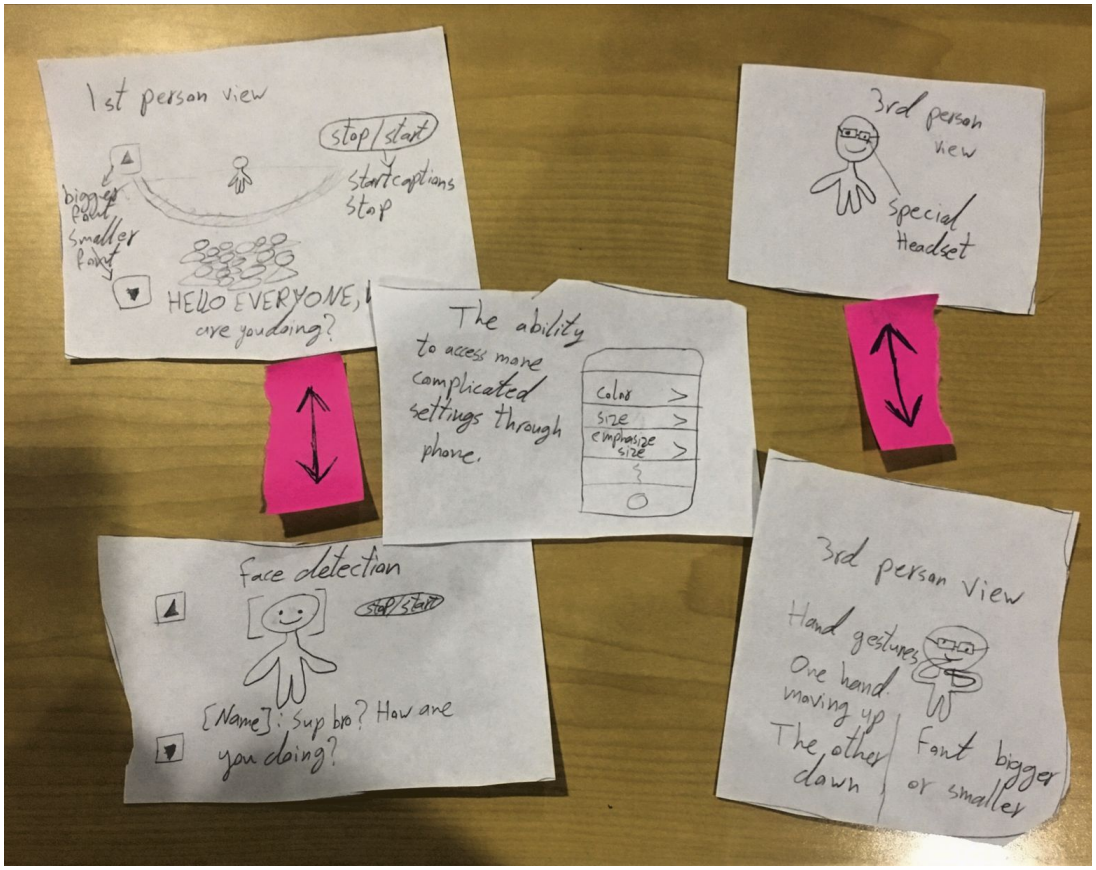
People who use text as the main way to communicate with others face a major issue. The issue is that text fails to display tone and emotions conveyed through verbal communication. For a deaf student, this poses a real problem; For example, such a student that cannot hear what is said relies on other means that don't allow him to be able to grasp what points the professor is emphasizing during a lecture. Our application aims to solve the problem in the following two settings in the following manners: The first is using live captioning along with manipulating the size and shape of the text to reflect how loud they are said (this can easily express points of emphasis in a lecture). The other approach is by enabling texting friends through also manipulating the text in a manner that will reflect one's emotions (changing color of the text, motion of the text, bold and underlined text, etc).

Sketches

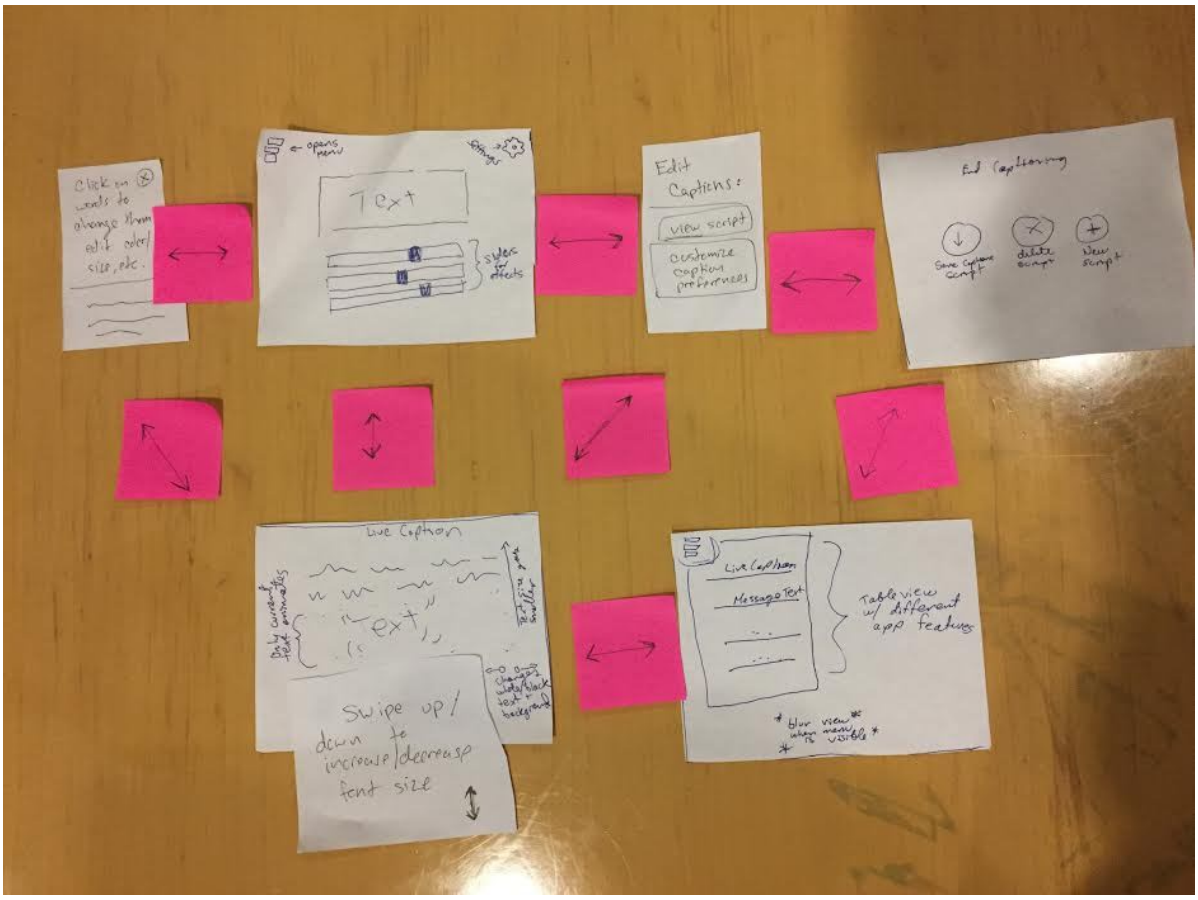
Concept Sketches / UI Sketches



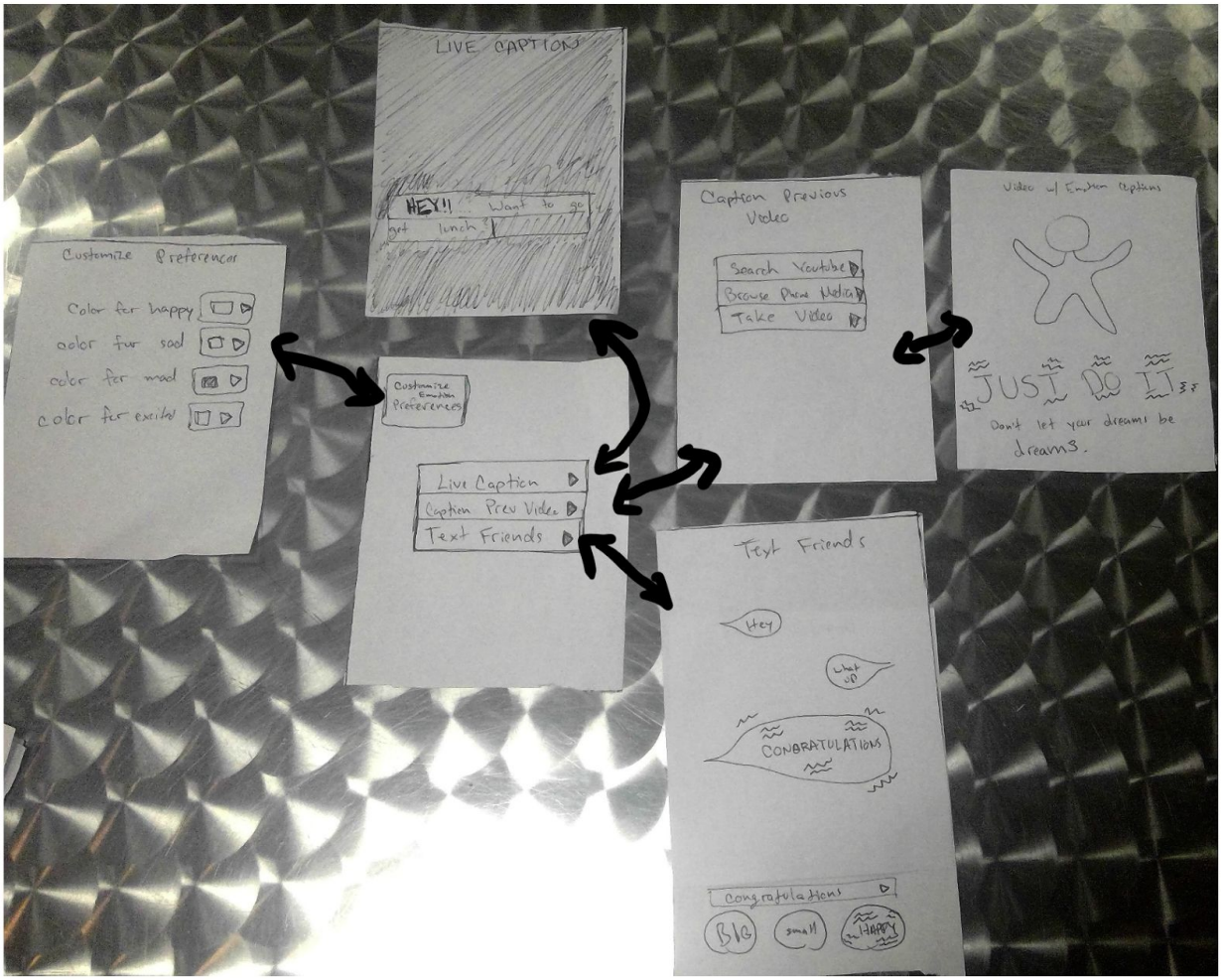
Sketch 1 -- Smart TV for watching shows + an interface for producers to add emotions.



Sketch 2 -- Augmented reality interface using glasses. Similarly live captions and manipulates text in real time.



Sketch 3 --iPad app- Interface incorporating live captioning along with live editing of captions.



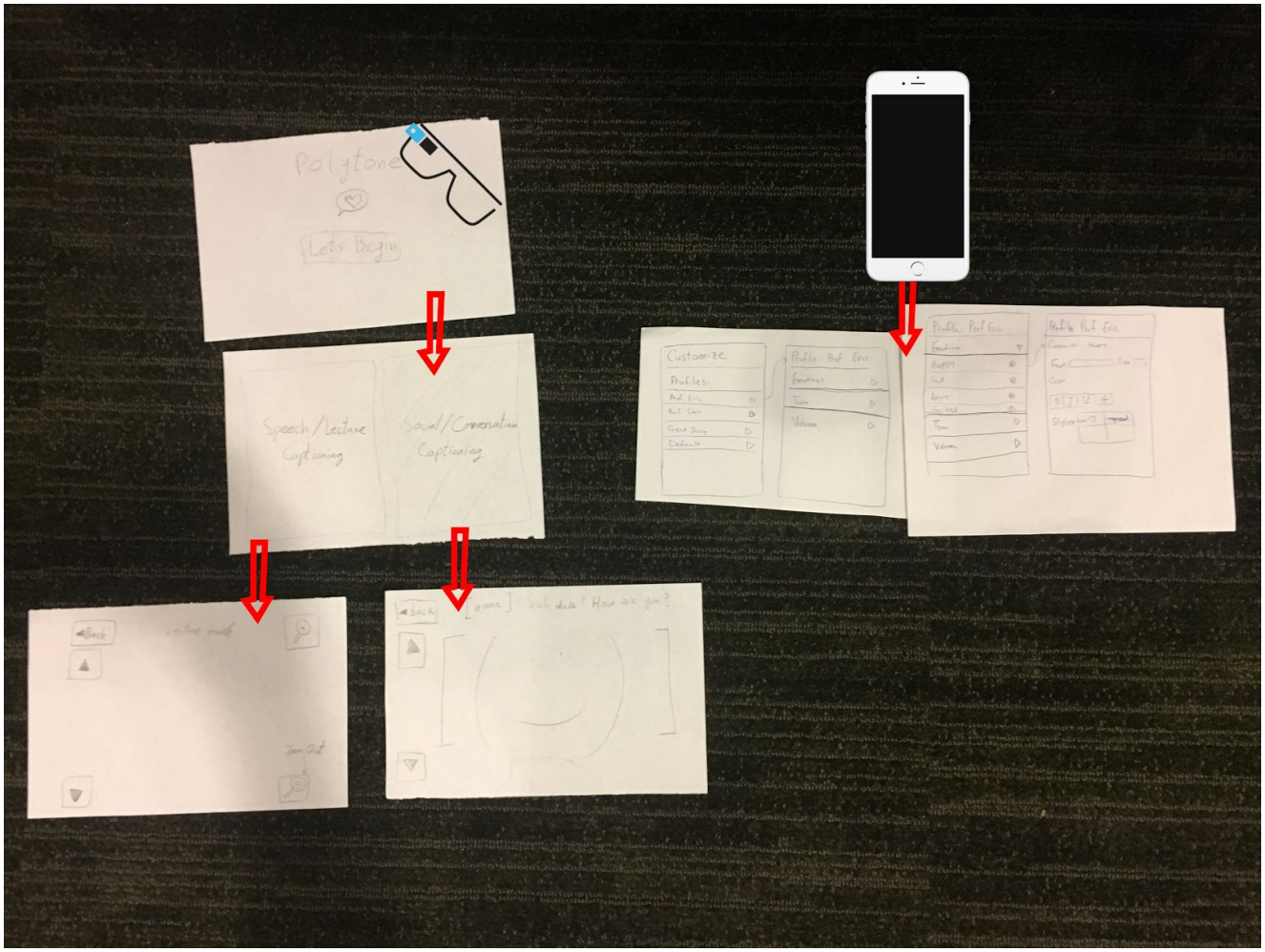
Sketch 4 --iPhone app- Interface incorporating live captioning, video captioning, and texting friends.

Selected Sketches:

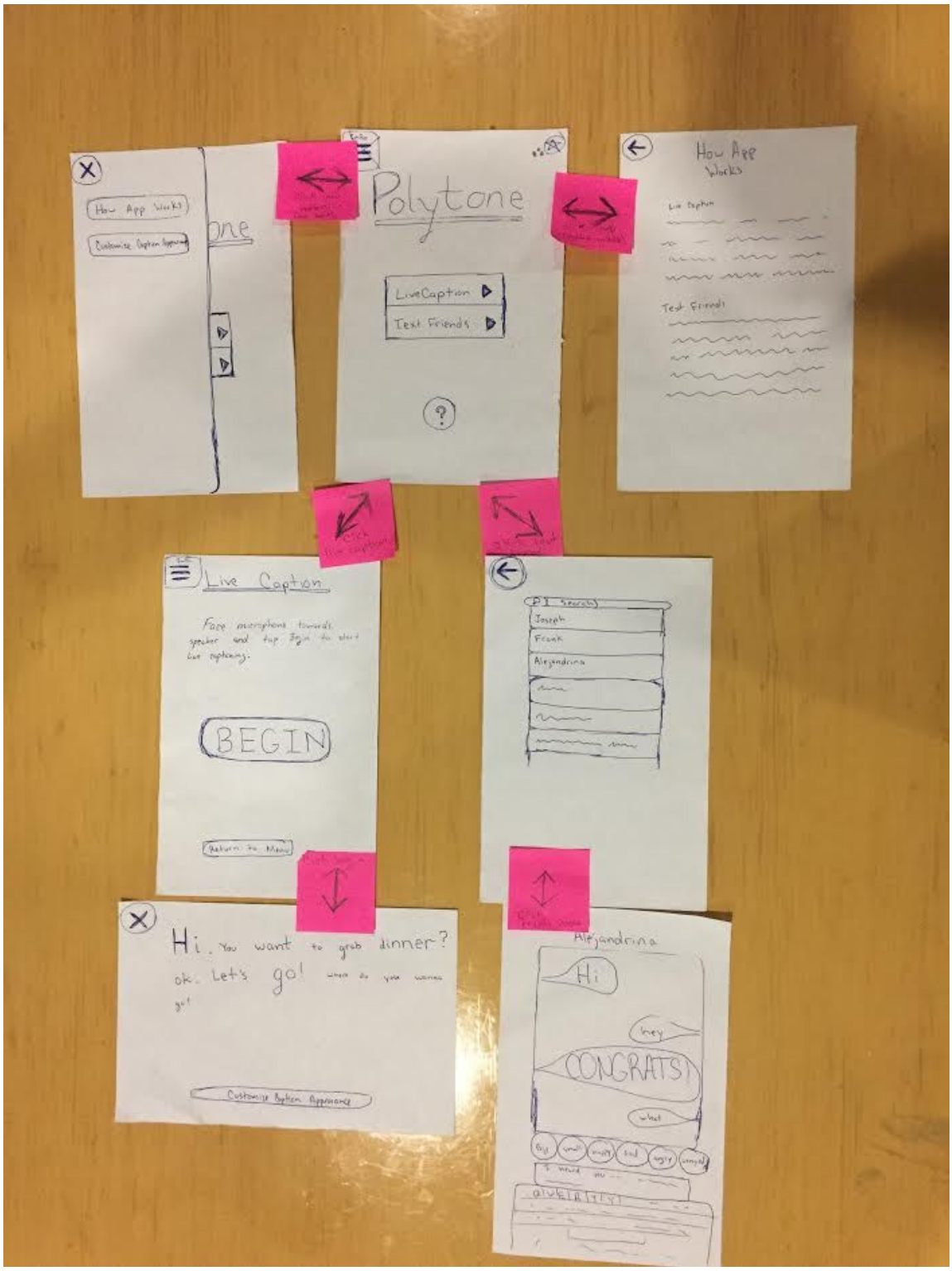
We selected our top 2 prototypes to be the Augmented Reality prototype; in addition to the iPad prototype.

The AR prototype provides a way for one to read emotion captions while also getting to watch the speaker/professor. Moreover, it enables users to use face detection to know who is speaking, and who might be the person talking to them. This enables another dimension for viewing emotions in text for the text appearance can change based off the individual you are communicating with.

On the other hand, the iPhone/iPad application allows users to send messages with emotion. More importantly, it gives users the ability to caption live conversations (useful for a deaf person trying to have a conversation with another that doesn't know sign language) and lectures as well. It also has the caption customization ability. Below are both two refined storyboards of the two prototypes.



Sketch 5 -- AR after refining

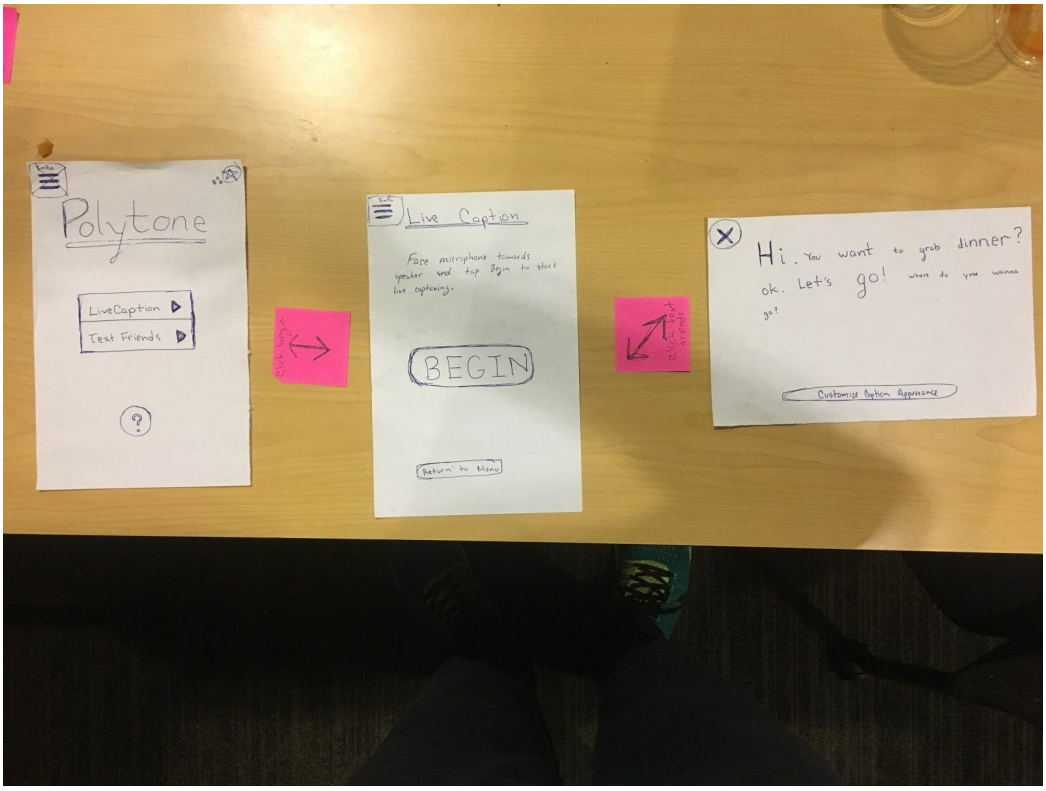


Sketch 6 -- iPad/iPhone after refining.

Selected Interface

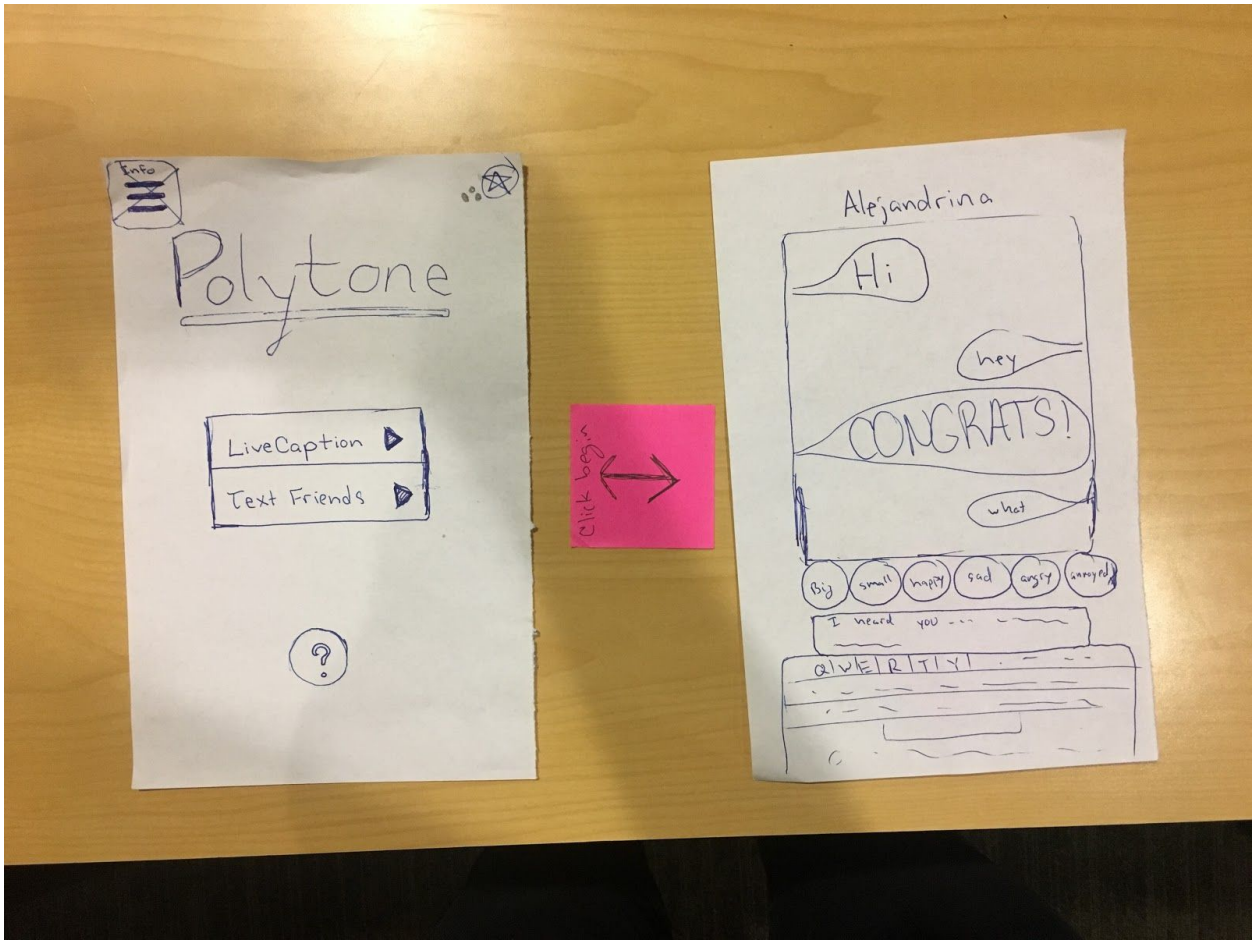
Storyboard for 3 tasks

Task #1 - As a deaf student, I want to be able to visually understand tone, style, and volume of a lecture.



Here, this is a simple task. All what the user has to do is clicking few buttons to reach the screen that starts captioning with style. While this is typically for lectures, it could be as well used in meetings, or 1-on-1 sessions. The problem to use it in such situations is that it doesn't have a direct way to reply.

Task #2 & Task #3: 1-on-1 conversation captioning & iMessage extension represented on the same storyboard.



While we were at this stage, our intuition was that if you want to talk to someone 1-on-1, then there's no need for a separate screen to do that other than messaging with style. This resulted to the fact that we have in this picture 2 tasks that could be done together. Any person can text my friends with style that reflects my emotions. In addition, in 1-on-1 live conversation situation the messaging could be done through bluetooth for example, so that you don't have to add the person you're chatting with. **Later** in our process, we revisit this decision, and we decide to do messaging and 1-on-1 in a completely different way and UI.

Reasoning for selection

The selected interface design had the best strong visual hierarchy from all prototypes by having a clear design. Because users have different approaches when using an app, the selected design had the best flow matching the flow of the work the user is trying to accomplish with each task in the app. We based our design on three main functionalities: the live captioning session screen, the customization of the app text, and the saving of the

captioning sessions. With this in mind, the clear design hierarchy with one primary action per screen keeps users in control by giving less of a cognitive load on the user.

	Pros	Cons
Prototype 1: Augmented Reality on Head-Mounted Display	- Minimal UI - Augments Reality - Easily accessible in glasses	- Stands out - Not widely available - Needs complementary UI on mobile
Prototype 2: Mobile application with messaging extension	- Accessible extension - Very easy to learn and use with its familiar mobile UI - More subtle as an assistive technology in class	- Doesn't greatly utilize augmented reality - Complicated integration of different forms of verbal communication all in one form of text.

Prototype

Our prototype has a main screen with access to three main features in the app. The home screen has access to the live captioning, customization, and saved live captioning sessions. There's an information button on the home screen as well, it shows basic info on how to use the app.

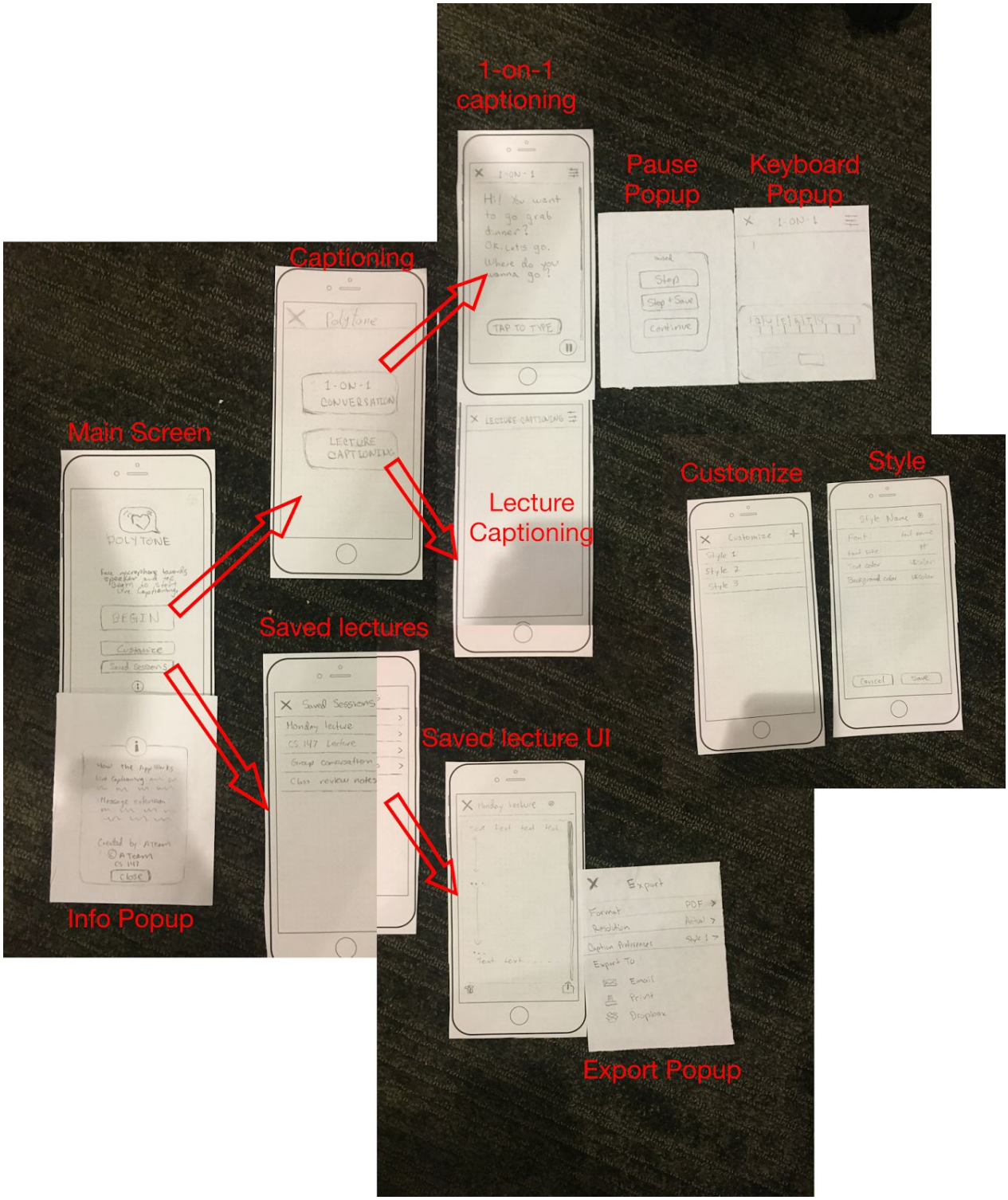
There are two main functions in the live captioning screen, and two of our tasks: 1-on-1 conversation and captioning lectures. On the live captioning screens, there's the option to pause/stop/continue/save the session. There's also immediate access to the customization screen in case the user wanted to change it while the live captioning session is going. For 1-on-1 conversations, there's an optional pop-up keyboard.

For customization, the user can save different styles and add/delete new ones. Add by tapping the (+) button on the top right. Swipe left on the table view cell to delete the style. The user can give a name to the custom style.

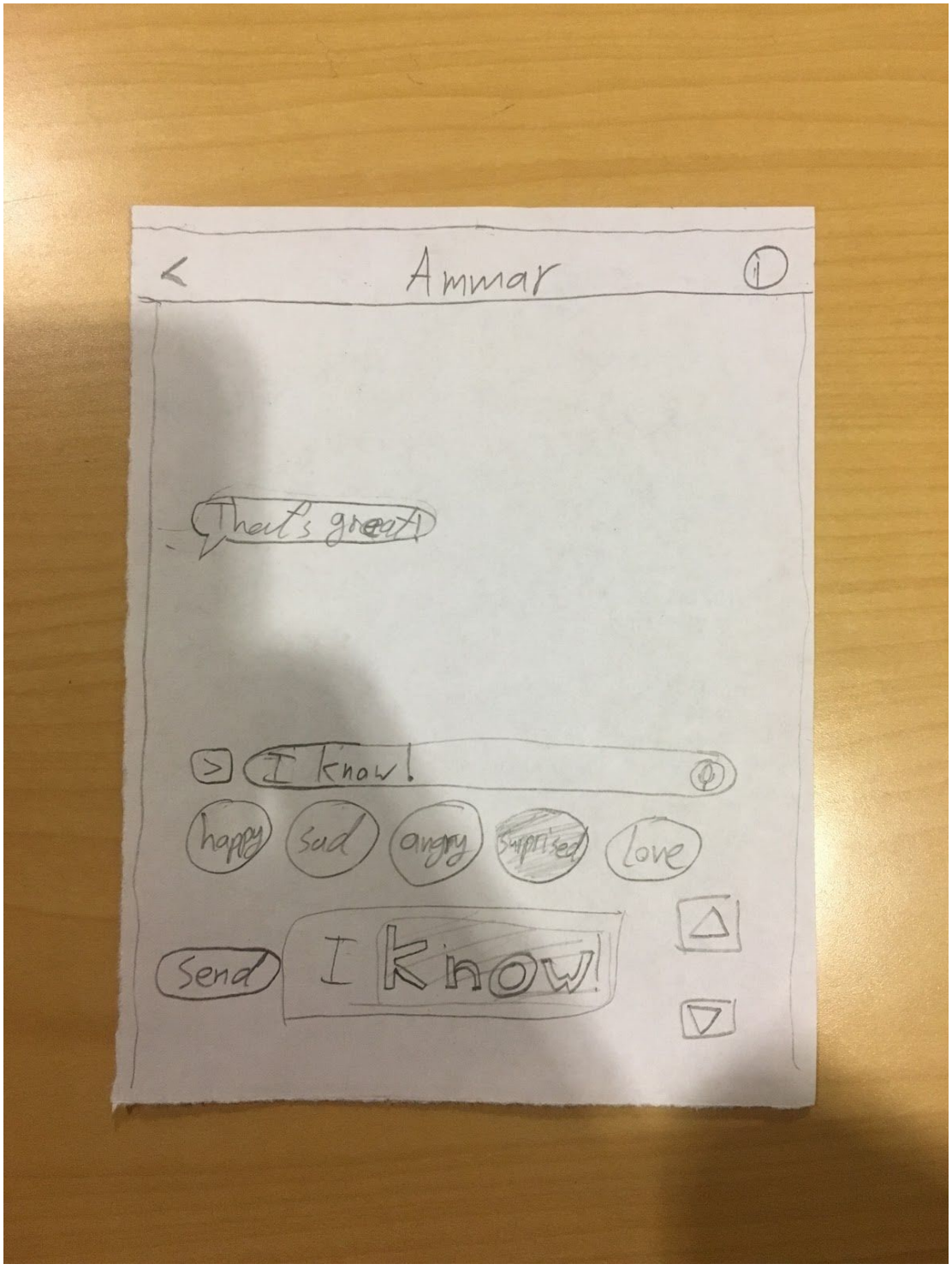
On the saved sessions screens, there's a table view with the customized names for each session. When the cell of the table view is tapped, another screen shows with the text of the sessions. On that screen, the user can view all the text, rename the title of the session, and share/export pop-up or delete the session.

In addition, we have a part that is an iMessage extension that allows users to stylize the text that they want to send to their friends. The idea is being that they give the extension

the text, and within the extension they add the emotions, size, etc. And when they send it to friends it gets sent as an image to so that the other person can see the style.



Lo-Fi Prototype -- Storyboard + Workflow



iMessage Extension Simple Example

Method

a. Participants

1. Student- Self identifies as deaf.
2. Student- Self identified to be hard of hearing.
3. Third person - Normal, no hearing loss.

We contacted local and campus hearing loss groups and organizations. We asked to be connected to people with hearing loss and were consequently put in touch with a variety of people. From here, we reached out to several people and were able to interview two with hearing loss. For our third interview, we found a friend of a friend who was willing to undergo the prototype testing process. The choices made sense because our app could be used by students of hearing loss, and normal students as well, but with a focus on students of hearing loss.

(We cannot disclose the identities of our participants due to the signing of a confidentiality agreement).

b. Environment

We performed all of the prototyping interviews in quiet, open, public and natural environments for each of our interviewees. Our first interview was performed on an empty table outside Old Union. Our second interview was performed in a nearly empty discussion room in the Markaz (where the interviewee was previously working on homework). We met our last interviewee in an open, nearly empty lobby of her place of residence. All 3 places were picked by the interviewees depending on their preferences.

c. Tasks

1. Use this app to live caption a lecture with emotion. Task one had simple complexity and a simple objective. We aimed to see how simple it was to navigate through the app along with how effective emotion enhanced text would be.
2. Use the app to have a 1-on-1 conversation with someone using emotion live captioning. This task was a more moderate task for it involved more steps for successful communication with another individual that was deaf (or for a deaf person to communicate with someone who is not deaf). We wished to assess the usability of a 1-on-1 chat feature between those with hearing loss and those without (when a communication barrier exists).
3. Communicate with a friend while fully displaying emotion only through text. Although the objective here is simple, it is a very complex task for it involves the added complications of

the variety of aspects of verbal communication and the attempt to convert them to a written format. We aimed to evaluate how difficult, feasible, and effective such a task would be.

The are further assignments we asked the user to accomplish while navigating through the app. They are not part of our three main tasks for the app. However, we asked the user to accomplish such tasks in order to evaluate how easy it would be to navigate through the app.

- Create a new caption appearance style.
- Record and save a lecture's emotion captions.
- Share your saved lecture with captions with a friend through email.
- Navigate to the help/app info screen.

d. Procedure

Firstly, we asked the interviewees a few general questions about their hearing loss and previous difficulties they experienced with lectures and casual conversations. We did so because we know that the range of people who suffer from hearing loss is vast. We wanted to know, for our references, if such an app would be usable to the people tested.

Next, we introduced our project and gave a brief overview of the problem and the proposed solution of manipulating text to display emotion. We then went to explain how we would carry out the prototyping interview process along with providing a quick demo. We explained how the interviewee should think out loud while navigating through the app.

Lastly, we showed them our idea of the iMessage extension, and asked them whether they would use it or not, and what emotions would they love to use the most. At the conclusion of the prototyping interview, we asked the interviewee if there was anything specifically that they really disliked or liked about the low-fi prototype app.

e. Test Measures

While we carried out the prototyping interviews, we tried to evaluate what the interviewees thought of the app concept, its usability, and its effectiveness in tackling the problem at hand. To do so effectively and consistently among our three interviews, we implemented the following testing measures:

- Only ask the user to carry out the task and restrict from aiding them in navigation through the app entirely.
- Repeat the task to the user if they are unclear of what it was or if they seem to be completely off task (in which case they probably forgot what the task was).
- Let the user know if they have not successfully completed the task.
- Give the user as much time as necessary to navigate to the designated window to complete the task (regardless of how long it may take).

- We kept track of which tasks took the longest and were the most cumbersome for the users.
- We made sure to keep track of when the users were confused about a certain button or if they thought there should be another button or feature to ease usability.

Results

We obtained a range of results regarding the app interface, missing buttons and features, the potential usefulness, and usability of the app. We divided our based off the interviewees. Included are the most important points from the opinions and experiences of each of the interviewees:

All our interviewees

- Believed there should be an evident and more accessible pause and stop button when live captioning a lecture.
- Found the interface to be very clean, clear, and easy to navigate.
- Thought that such an app could be useful and that they would use it if it existed.
- Thought that the iMessage extension was a cool idea, and they would use it for emotions like surprise, happiness or love.

Interviewee 1

- Found the interface to be very clean, clear, and easy to navigate.
- Expressed concerns with how accurate current speech recognition software would be in correctly and accurately recognizing what is said.

Interviewee 2

- Expressed concern with the fact that the live captions for lectures exited the screen too quickly and that it would be better if they stayed on the screen and the windows simply continuously scrolled down.
- Expressed concern regarding the usefulness of the separation of lecture captioning and 1-on-1 conversation captioning in the app (since she did not find the adding functionality of typing to communicate with another person particularly useful).

Interviewee 3

- Found it cumbersome to navigate back to the start menu when attempting to accomplish a different task and wished there existed a button that directly took you to the home page.
- Believed it should necessary to share a saved lecture captions directly from the screen you are live captioning it from instead of having to navigate back to the home page.

Discussion

The results from our interviews gave us insights into the overall concept and underlying structure of our app, as well as into the various small interface choices in our app screens.

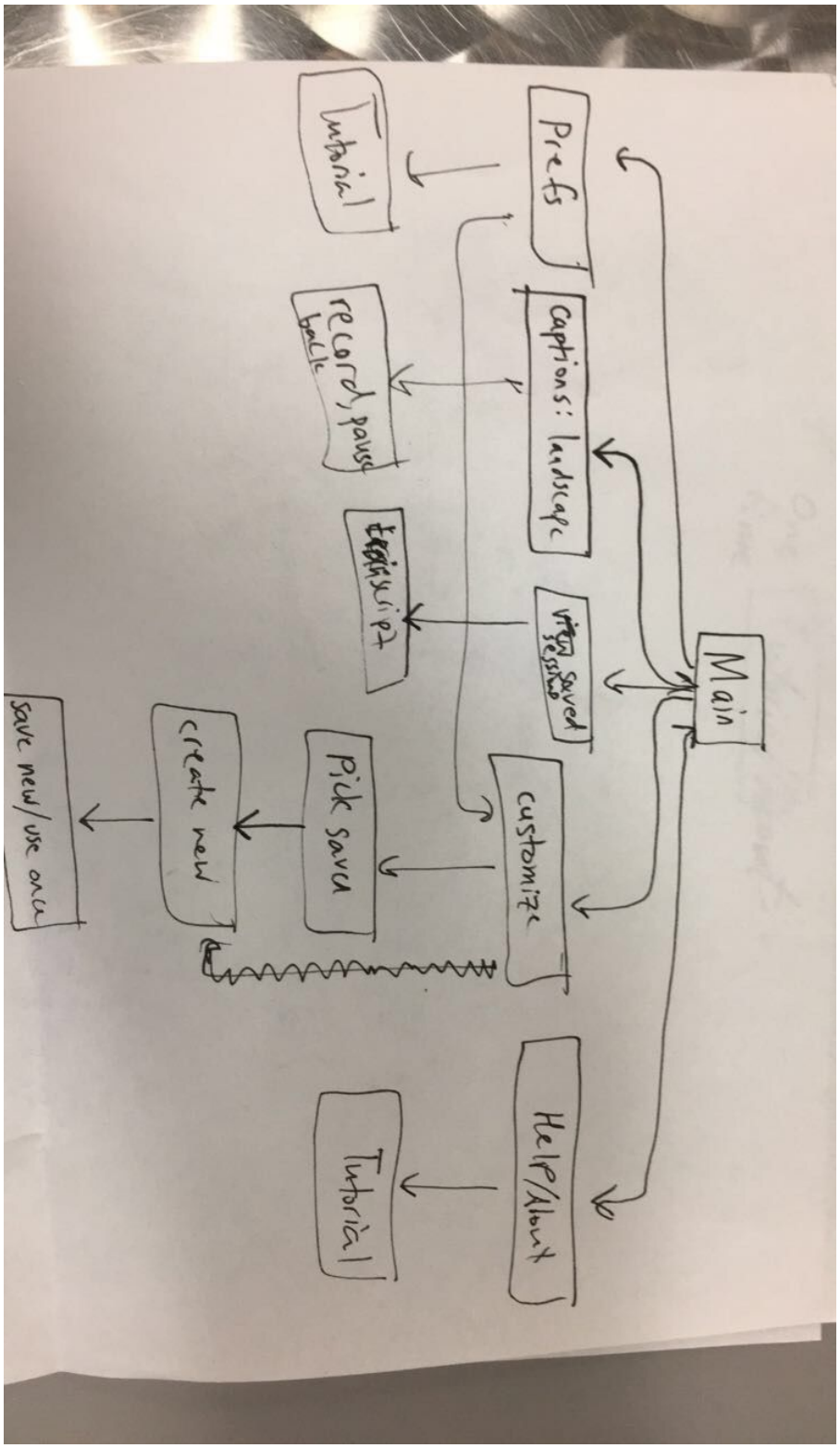
Overall, it seems that our app was well-received both from users who identify as hard-of-hearing (HOH), who are our target demographic, and, surprisingly, users who are not HOH. HOH students recognized the element of engagement the application can bring to captions to replace the monotonicity of the current captioning system. They also found use for it outside of the setting we envisioned, the classroom, to contexts such as watching movies, plays, or musicals. For non-HOH students, however, the main appeal of the app was the fact that it was tailored specifically for students and allowed them to accomplish tasks such as recording, saving, and exporting lectures easily in a textual format.

One of our main results made us reconsider our division of the app into two main tasks with different flows -- captioning lectures and captioning conversations -- to perhaps combining them into one interface that is versatile enough to accomplish both tasks. The users found it confusing that the conversation task involves continuous and persistent captions in portrait while the captioning task had momentary subtitle-style captions in a landscape screen. We found a way to combine the tasks to make them more intuitive would be to have them have the same interface of a continuous, scrollable feed of captions with the phone being either landscape or portrait.

The results also made us aware of some essential elements that were missing from the interface such as: better controls for recording, pausing, playing, and saving while live captioning; confirmation prompts for different tasks; the ability to delete saved styles in addition to adding them; and the ability to share/export captions immediately. Moving forward, our results will make us simplify some key aspects of our design to make our app more intuitive and cohesive to accomplish our main tasks, and to pay attention to interface choices that make the app easy and comfortable to navigate.

Some limitations of our prototype experiment, however, are in the actual accuracy and speed of speech recognition and whether these technological limitations will hinder the user experience. The experiment also didn't test the aesthetics of our text styling, which is an important part of our application; but that is because we focused more on the concept rather than the actual implementation -- the users, however, did like our initial representation of the test and thought it did enhance their reading experience.

Appendix



Workflow

Consent Form

The **ATeam** application is being produced as part of the coursework for Computer Science course CS 147 at Stanford University. Participants in experimental evaluation of the application provide data that is used to evaluate and modify the interface of **ATeam**. Data will be collected by interview, observation and questionnaire.

Participation in this experiment is voluntary. Participants may withdraw themselves and their data at any time without fear of consequences. Concerns about the experiment may be discussed with the researchers (Abdallah AbuHashem, Ammar Alqatari, Ahmed Shuaibi, Alejandrina Reyes) or with Professor James Landay, the instructor of CS 147:

James A. Landay
CS Department
Stanford University
650-498-8215
landay at cs.stanford.edu

Participant anonymity will be provided by the separate storage of names from data. Data will only be identified by participant number. No identifying information about the participants will be available to anyone except the studentresearchers and their supervisors/teaching staff.

I hereby acknowledge that I have been given an opportunity to ask questions about the nature of the experiment and my participation in it. I give my consent to have data collected on my behavior and opinions in relation to the [TEAM NAME HERE] experiment. I also give permission for images/video of me using the application to be used in presentations or publications as long as I am not personally identifiable in the images/video. I understand I may withdraw my permission at any time

Name _____

Participant Number _____

Date _____

Signature_____

Witness name _____

Witness signature_____