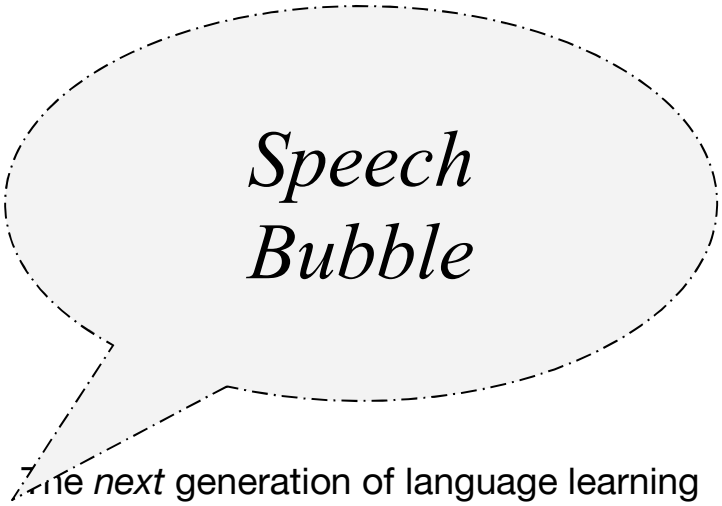


The *next* generation of language learning and speech suggestion wearables

Amrutha Gujjar
Arielle Menn
Mary Edwards
Sachin Bharati

In today's globalized world, we meet people from all over the world, people from many different countries who speak many different languages. Between accents and colloquial phrases and slang, it's all too easy to become *Lost In Translation*. We aim to tackle the problem of communication difficulties head on, using our 3-pronged cross platform approach. By integrating a small, earpiece wearable, a smart watch interface, as well as a smartphone application, we are able to help people feel more at ease during their conversa tions by providing them aid along the way. The earpiece is able to detect parts of a conversation that may be poorly phrased or difficult to understand, using our sophisticated natural language processing algorithms. We create a consistent feedback loop by giving more consistent updates to the watch interface, as well as more detailed reports to the smartphone application. By providing the client with regular, reliable and honest feedback, we are able to help them track their progress in achieving greater conversational fluency.

Design Research Goals, Stakeholders, and Participants:

Throughout our design process, our primary audience focused broadly on people who were looking to improve their conversational English skills. Initially, our design research spanned anyone from business professionals who were interested in achieving a more professional tone in the workplace, to students who need to incorporate more advanced vocabulary into their academic projects and papers. Our initial design research targeted participants who were young business professionals in fields such as marketing, and sales, given the communication driven aspect of these professions. Given professionals in these fields typically face many day to day encounters and difficulties in speech, it made sense to understand the kinds of communication problems they encounter. We ultimately decided that our stakeholders are generally non-native English speakers who want to achieve more fluency and clarity in their English skills.

Our primary research approach was to conduct interviews. This was so we could focus on exploring and learning about our various target groups and participants, and how they felt about the problem we were trying to address. Our focus throughout the interview process was to better understand how non-native English speakers used conversational english in their everyday lives, whether that be in the realm of social, or professional, or educational settings. We thought this was a valuable component of our design research, because it gave great insight into the situations in which they felt their English skills were not sufficient, and what would be helpful to them in these situations. Our tracking device would be able to fill that gap by better pinpointing improvement areas. In our interviews we talked to students who grew up in different countries and had difficulties communicating with teachers and peers, we talked to engineers and employees who felt that small talk with coworkers was a challenge, especially in the case of colloquial jokes and obscure references. This aspect of our design research gave us insight into how this broad group of people all use English in various social contexts, and what we found was that the feedback was something that we needed to provide rather immediately. Though the long term tracking is useful to show progress and improvement, the small nudges to make certain pronunciations more clear or to explain certain concepts should be administered immediately. That's why our design research was more geared towards constant and immediate feedback, rather than consistent tracking in a diary study. Additionally, we moved away from self-reporting, since as a non-native English speaker, sometimes it can

be difficult to acknowledge or realize when something being said is unclear. We figured some kind of objective third party listening device would give us the most candid results.

Design Research Results and Themes:

In our design research, we found that each subject had a fairly unique perspective and set of problem areas they considered worth tackling, but when we asked subjects to reflect on suggestions we'd gotten from other interviewees, we started seeing patterns with more popular complaints and suggestions.

One theme that emerged from our interviews was a frustration with finding the proper word or idiomatic phrase in the moment. People described the frustrating experience of having a specific word they want to say but not remembering the vocabulary, and needing to describe the concept or object in a roundabout way until the other person guessed what they meant. In many cases these words were fairly new, just heard from a coworker or classmate the previous week. Similarly, understanding idiomatic phrases (e.g. "let the cat out of the bag") and learning when to use them is a big challenge for several of our interviewees. Beyond understanding what is meant when someone else uses a phrase like that, it's very difficult to have the thought to use a metaphor when they themselves are speaking.

Another complaint that was commonly cited was difficulty in understanding quick, slurred, or accented speech. One participant mentioned that listening to people with certain accents was more difficult than hearing people with other accents, and that quick or badly enunciated speech also was challenging; after that interview, the other people we interviewed quickly and enthusiastically agreed when we brought up the thought with them. Several added that they need to learn to cope with having missed bits of class material or information at a meeting because they simply did not understand what was said in that moment, and the speech moved too fast or was too public to ask the speaker to backtrack and slow down.

One habit that became clear among our interviewees was the desire to be corrected, albeit in a friendly manner. All of our participants mentioned that they not only want to improve their English, but find it helpful when others give them gentle suggestions or corrections. One popular comment was that though occasionally those corrections can be annoying, for the most part they are not said in a condescending way and the person takes them well. People

we interviewed did not like to be constantly nagged about one detail of their speech, and appreciated new suggestions and gentle corrections.

Other comments came from one or maybe two subjects throughout our research, but were not agreed upon by the other people to be important areas of struggle or frustration. These included the frustration around others mispronouncing a name, anxiety surrounding professionalism in written communications like emails, and confusion when reading emails and other written conversations where intonation does not hint at the mood or intention behind the sentence. Though these were valid complaints from individual participants, most other people in our research group did not respond strongly to those sentiments when asked, so we did not consider them as seriously when we went into our task definition phase.

Answers to Task Analysis Questions:

Task 1: Slowing down speech so it is easier to understand:

- **Who is going to use the design?** The target demographic we have in mind is people who are not native english speakers and may have a hard time understand area specific and culture specific vernacular
- **What tasks do they now perform?** The tasks will act as the bridge between a non traditionally used phrase or word and its actual meaning.
- **What tasks are desired?** The task we desire is the ability for the application to record certain parts of a conversations that people feel unsure or uncomfortable with with an end goal of being able to giving them the ability to better understand what was originally unclear with.
- **How are the tasks learned?** The tasks would be learned by people who have a desire to be able to understand the conversations they have better, as a backup. And then spread by word of mouth from successful usage.
- **Where are the tasks performed?** The tasks are performed during conversations or one way presentations such as a business meeting, a talk or even a lecture.
- **What is the relationship between the person and data?** The person provided the date through the watch/earpiece while they are having a conversation with someone

else or when listening to a lecture/presentation. Anywhere where there is a chance for verbal misunderstanding from mone party to the user of the product.

- ➔ **What other tools does the person have?** The person will have an earpiece and a watch they can use to make recording and playback as easy as possibly can be. Having these two devices will make recording and playing back information much more seamless than having to use once phone every time someone wants to record what they they aren't sure about and then play it back with their phone if it is during the conversation.
- ➔ **How do people communicate with each other?** Through conversation with others or other situations that require the client to have to listen to someone else talk
- ➔ **How often are the tasks performed?** The tasks are performed whenever someone is listening to someone else and can be performed after the interaction to view previous clips of conversations that they did not understand.
- ➔ **What are the time constraints on the tasks?** This part depends on the user as they have to know what is confusing them and what part of conversation to record. After this they have the luxury to replay and understand old confusing conversations at their convenience
- ➔ **What happens when things go wrong?** The application could interfere with sometimes conversation adding a barrier between the parties involved (think google glass) or the application could misunderstand the data being input to it and give incorrect or misleading information

Task 2: Remembering how to use common phrases:

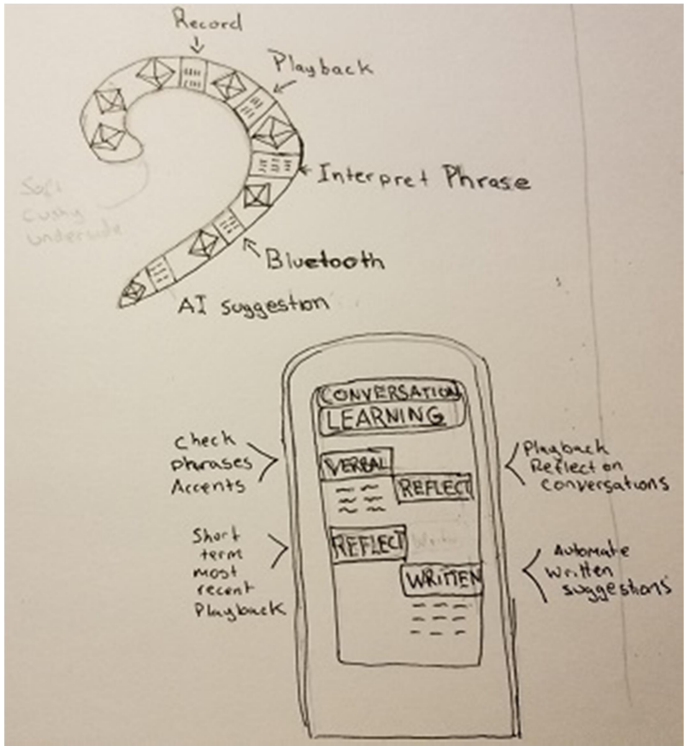
- ➔ **Who is going to use the design?** Anyone who at some point struggles to remember the meaning of phrases.
- ➔ **What tasks do they now perform?** They act as a resource users can use to understand vernacular they hadn't known before.
- ➔ **What tasks are desired?** For a person to be able to gain knowledge of local speech that they were unfamiliar with before.
- ➔ **How are the tasks learned?** The tasks are leant by user input updating the library of local or regional vernacular to a pre-existing library of words.
- ➔ **Where are the tasks performed?** The tasks are performed on the watch's interface, with either the watch or earpiece outputting the results of the input.

- ➔ **What is the relationship between the person and data?** The person inputs the data into the watch to produce the results of the phrase they requested or the correct pronunciation of a word they wanted to know about.
- ➔ **What other tools does the person have?** The tools essential for the person again are the watch and the earpiece so they can enter what they desire to learn about that they are unaware of and for the earpiece to explain or give the pronunciation for a word unknown to them.
- ➔ **How do people communicate with each other?** There is not much communication with others with this application of the product as people will mostly be entering data that they desire help to better understand.
- ➔ **How often are the tasks performed?** The tasks are performed at the user's discretion whenever he runs into something that the they feel as though they need help understanding.
- ➔ **What are the time constraints on the tasks?** There aren't any it is upto the user when he or she wants to request information from the application.
- ➔ **What happens when things go wrong?** Issues can arise when the user mis-remembers something they were unfamiliar with or incorrectly put input into the application that could give results for something the user didn't actually want. It is also possible for the application to return incorrect data to what the user actually desired.

Proposed Design Sketches - "3x4"

The idea of design1 is a wearable technology, a decorative earpiece that can either connect to an application or be used as a stand alone device. It will provide audio recording, playback, common phrase checking in addition to interpretation of heavily accented words.**Task 1: Pronouncing words correctly in a meeting for someone with an accent, (as it could playback speech) Task 2: Slowing down what somebody is saying so they are easier to understand** Furthermore there is also an Artificial intelligence option that provides conversational suggestions.**Task 3 Remembering how to use common phrases.** Additionally, the earpiece connects to a media device via bluetooth to allow for automated formal email templates(**Task 4: Writing an email to a boss asking for a raise)**, and or access to data to further

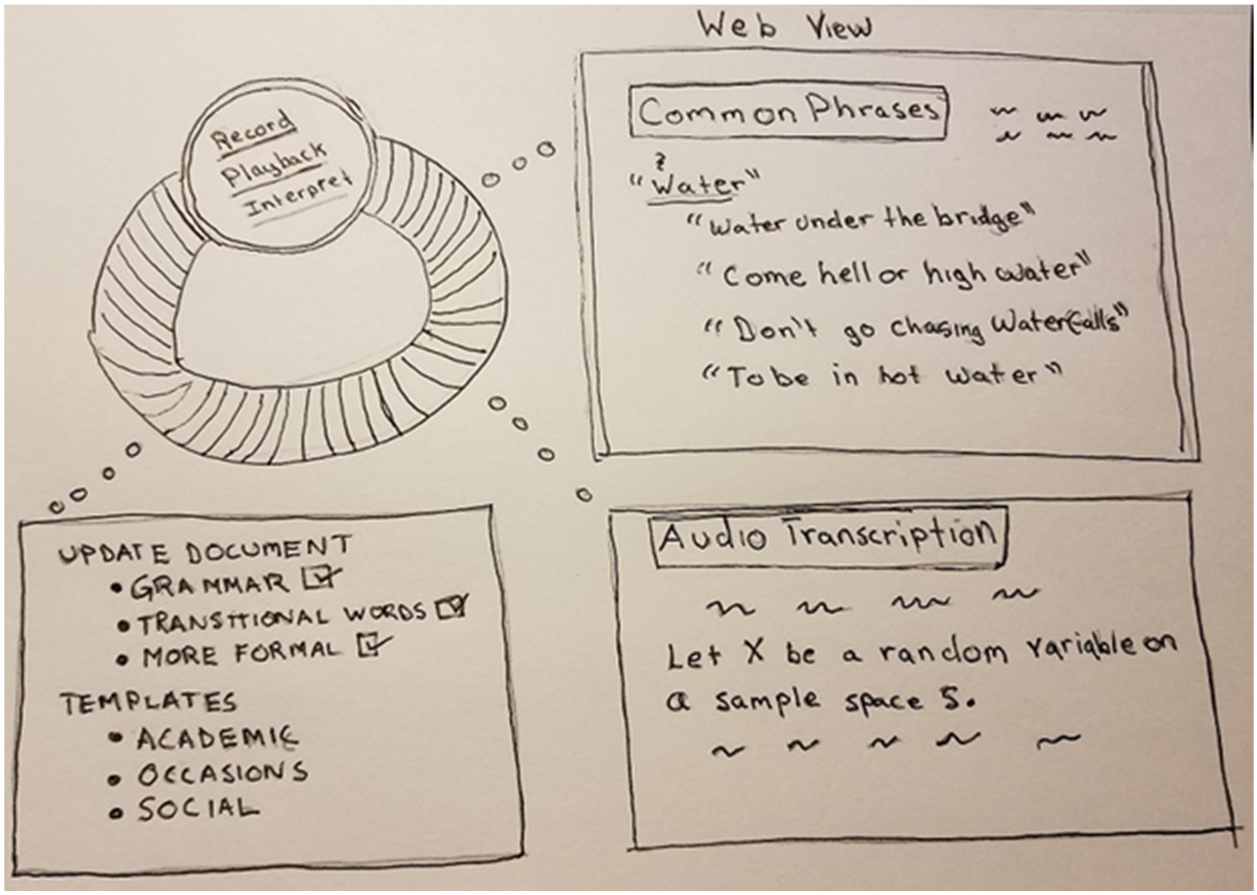
analyze/reflect on previously held conversations.



Design 2: Wearable Smart Watch with Conversational learning Application (also allows for ambient listening).

The idea of this design is wearable technology, very similar to design 1, except instead of a decorative earpiece, this is a fashionable watch, with more features. It has a touch screen which allows for easier access to its features; providing audio recording, playback, common phrase checking, and interpretation of heavily accented words. Furthermore there is an Artificial intelligence option that provides conversational suggestions, as well as ambient listening. Additionally, just like the earpiece, the watch connects to a media device via bluetooth to allow for automated formal email templates, and or access to data to further analyze/reflect on previously held conversations. Therefore it supported the same tasks as design 1: **Task 1: meeting for someone with an accent, Task 2: Slowing down what somebody is saying so they are easier to understand.Task 3: Remembering how to use common phrases. Task 4: Writing an email to a boss asking for a raise. However, it also provided more user comfort,**

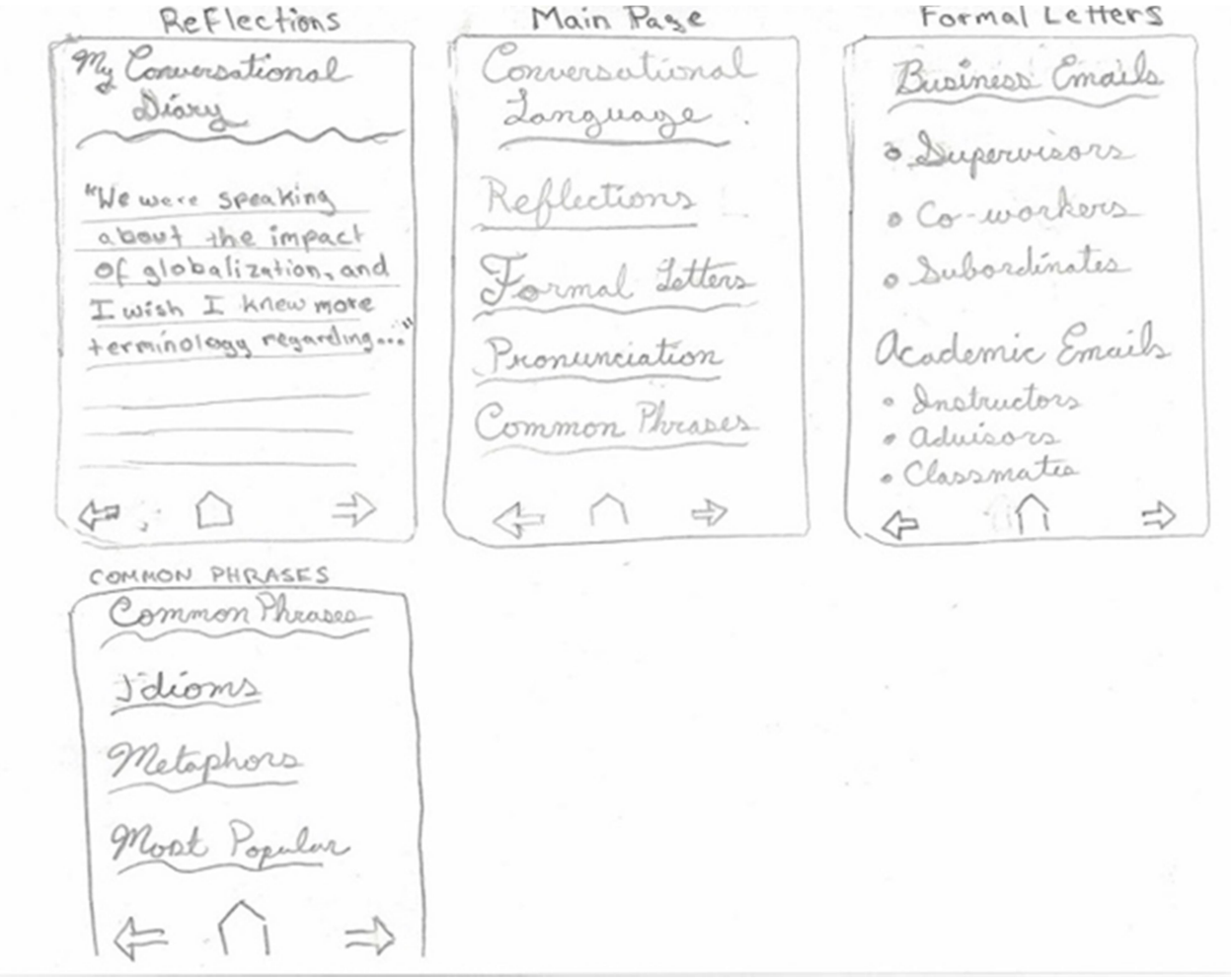
accessibility/ease of use.



Design 3: Web application/Mobile Application

This design is a simple web application/mobile application that allows the user to interact at a leisurely relaxed pace. It provides supportive features such as a reflection diary storing previous conversations, focused information on pronunciation, common phrases, and writing formal letters. With the reflection diary it supports Task 1: Wishing after a conversation that you had known or remembered a word to better describe yourself. The three specific categories are pronunciation, common phrases, and writing formal letters, each having their own distinct sections which support catering to Task 2: Remembering how to use common phrases, Task 3: Pronouncing words correctly in a meeting for someone with an accent, and Task 4: Editing a paper or professional document for more formal English, respectively. It is meant to have a more eloquent tone to it and that is why cursive is chosen as a font style. With an air of sophistication, classical music may be considered to be played in the background while one uses this app. Allows for searching and practicing of common phrases, providing access to a well explained knowledge

base of common phrases, idioms, metaphors, and most popular phrases as well.



We chose to focus on wearable technology integrated with a web/smartphone application. The design chosen was a merging of both the wearable watch/application design and earpiece/application design. Since feedback from the critics suggested people would prefer a more inconspicuous wearable; pushing buttons on an earpiece, or an earpiece that stood out was not wanted. However, we continued to like the support of being able to hear aural data in real time. Having an application built for any pre-existing smart watch could easily be adapted to a phone or web application and would support our tasks. We believe our target users will be those who are not as comfortable with conversational English, as well as early adopters of wearable technology. By integrating our product with existing smartwatches, we are able to cater to our target users by making the technology both useful and accessible. Additionally, since the people using our product will want to be fairly discrete with regards to their tracking, so as to not interfere with the conversations they are having, our design brings together various technologies, such as the device, the application, as well as a notification system to bring the best feedback while augmenting instead of interfering with the user's conversations.

The tasks we chose were slowing down speech so it is easier to understand, and remembering how to use common phrases**.** These tasks were more compelling than the others for a few reasons. In our feedback during critiques, many people responded more positively and were more vocal towards the parsing of conversational data. We felt more drawn toward these tasks because we felt that help with, and tracking of, verbal colloquial interactions was a much less well-addressed problem space than that of written English. Other tasks like drafting professional or grammatical emails were interesting and prevalent, but had already been addressed with email clients and other products, so we wanted to explore this new problem space. A third reason to pursue these tasks was that we saw more potential for creativity in design when dealing with issues that arise in social and conversational contexts, since the experience of the person using it is seriously affected by the social norms dictating the interaction and the most direct solutions (like constant conversation recording or real-time feedback that interrupts an interaction) wouldn't work. Overall we felt more excited to tackle the challenges that these tasks presented. We thought that our design and task choice worked the best as it tapped into a side of conversational and communication aide that hasn't been tapped into yet and that we would be able to make the greatest difference in this area. We also liked this idea of integrations as it provides an interface to help people converse versus producing a separate hardware product to that's sole purpose is to aid communication and conversations.

Written Scenarios - "1x2": (1 page):

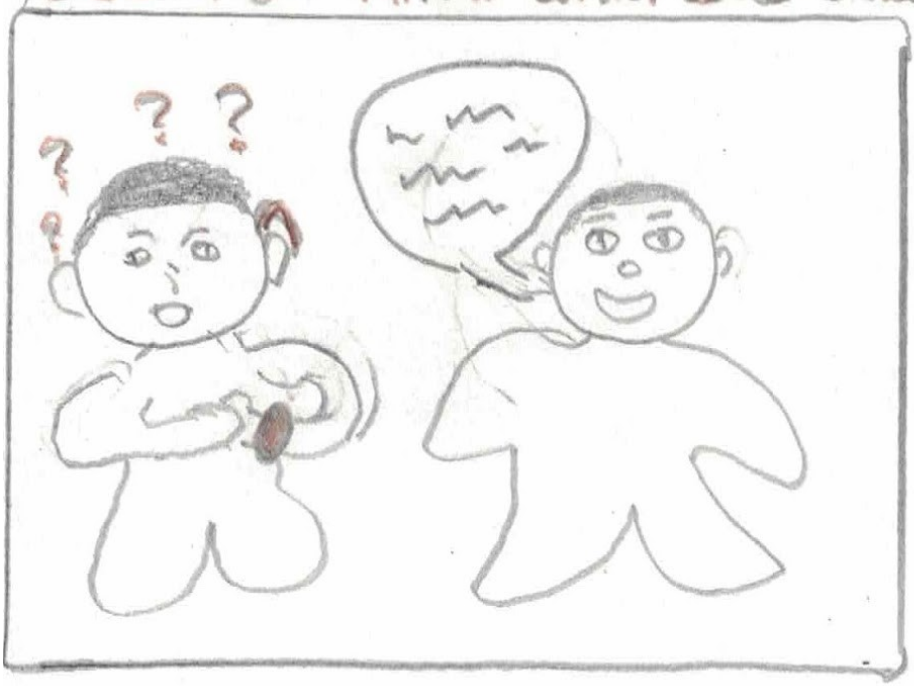
Pam is going to an important meeting where his boss's colleague from a neighboring branch will be presenting vital market research for Pam's upcoming project. Pam brings a notebook and pen to take notes. In the meeting, the presenter has very little text on the slide and seems very relaxed in front of the audience. He tells stories and shares information, but he talks fairly quickly and has what sounds like a French accent to Pam. Pam can't understand some of the phrases the presenter uses and feels like he's missing vital information for his project, in his

notes. To mitigate this, Pam opens the Speech Bubble app on his smartwatch. Whenever the presenter says something Pam doesn't catch, he taps the capture button and jots down the timestamp of the recording in the margin of his notes. After the meeting, Pam goes back to his office. He scrolls through the recordings on his phone, which has picked them up from his watch through an internet connection, and plays them one by one, and as he does the presenter's voice in the recording is slowed down. The phone screen displays a partial version of the sound bite, with blanks in the places of unrecognized words. Pam can use these two tools to fill in the blanks in his notes and gain a better understanding of the presenter's results.

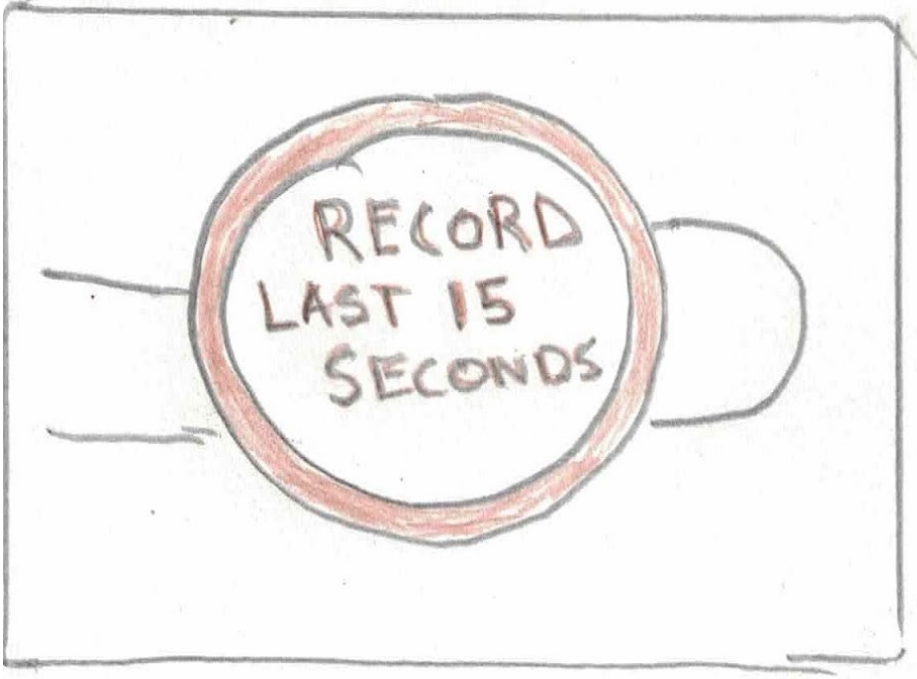
Jamie is chatting with a friend about the bad first date she went on last week. She explains to her friend, "the guy was just so...how do you say it? When someone thinks they are the most important person?". Jamie's friend helpfully offers the word "self-absorbed", and Jamie accepts the suggestion and continues with her story. "And he kept making fun of me for not knowing these random English phrases. Like he said something about a cat and a secret, I looked confused, and then he used the same phrase four more times to make fun of me for not understanding. I have it saved now though!" Jamie shows her friend her smartwatch. Pulling up the Speech Bubble interface, she speaks to the watch, saying "what is that phrase with the cat that I heard last week?" The watch speaker responds verbally as the screen lights up with the phrase: "to let the cat out of the bag". The display on the watch scrolls down to show "definition: to reveal a secret accidentally". Jamie explains to her friend, "if I had a second date with him I could use this to make sure he doesn't pull the rug over my eyes anymore--is that how I use that phrase?" She looks down at her watch to see it respond, "pulling a rug over someone's eyes means to hide some truth from them". "Oh," says Jamie, "I guess that wasn't really what I meant. Either way, he can't make fun of me anymore because I turned him down."

Storyboards of the Selected Design: StoryBoard One Slowing Down Speech For Understanding:

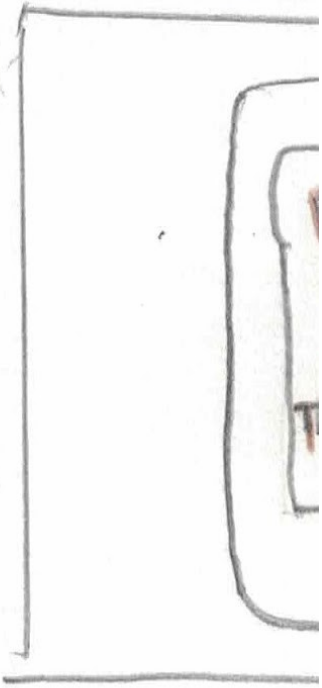
Does not know what was said.



Records for Review



Uses Ap



Storyboard 2 Remembering Common Phrases