

SpeakEasy

Integrating context-sensitive visual and aural language acquisition into your daily life.

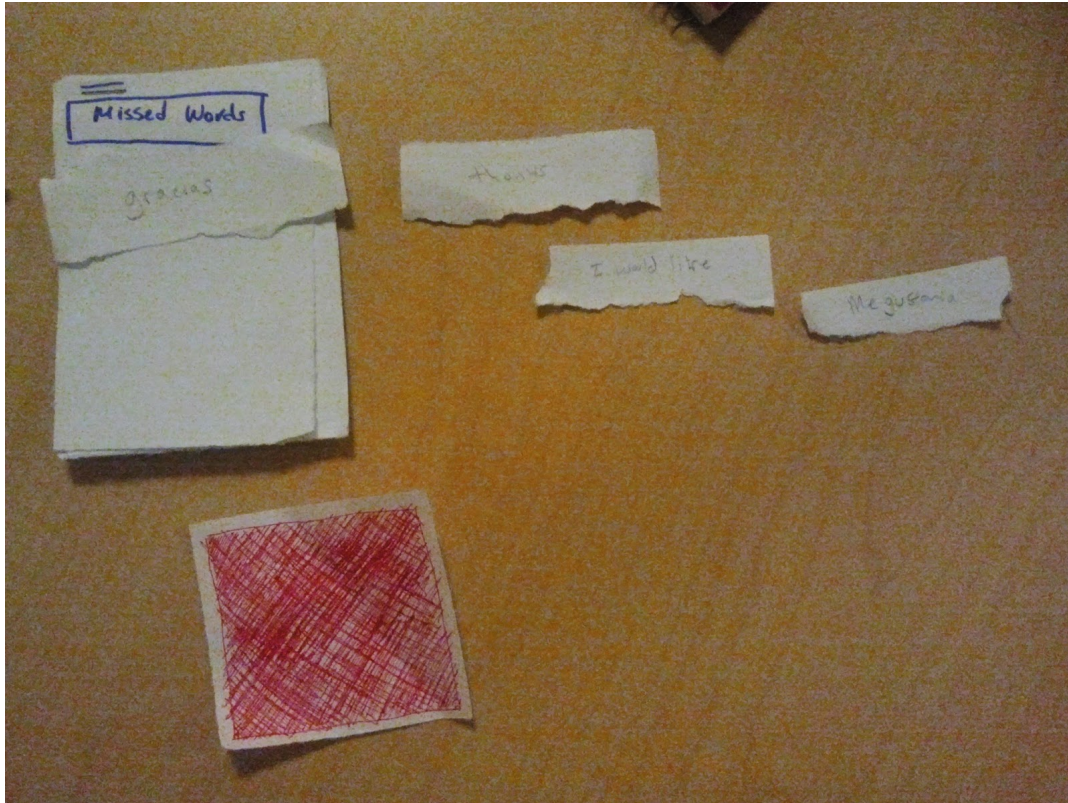
Gabriela Groth (Manager, Design), Eric Chew (Development), Tommy Truong (Documentation), Carlos Araujo (User testing)

MISSION STATEMENT

Language learning today is often removed from real world situations and partitioned into sizeable chunks that don't integrate well into a busy life. SpeakEasy strives to overcome these limitations by creating a context sensitive and on-the-go language learning environment. With instant speech suggestions, dynamic feedback loops, and helpful growth tracking, SpeakEasy allows users to learn and master a language even while performing other activities.

PROTOTYPE DESCRIPTION

Our prototype consisted of both a paper prototype of the mobile application as well as a half paper half mobile application prototype of the smart eyewear interface. The paper prototype was used to simulate the active pronunciation correction and integrated feedback functionality of our application. This paper prototype consisted of a basic mobile interface that allowed users to browse through their missed words as well as a red paper square that was displayed as part of the smart eyewear interface. During conversation, the red square was shown to users to indicate when they mispronounced a word. These mispronounced words were recorded and displayed on the paper mobile app, which users could interact with after conversation. The prototype allowed users to click on each mispronounced word to hear its correct pronunciation, swipe right to see a translation of the word, and swipe left to remove the word from their missed word bank.

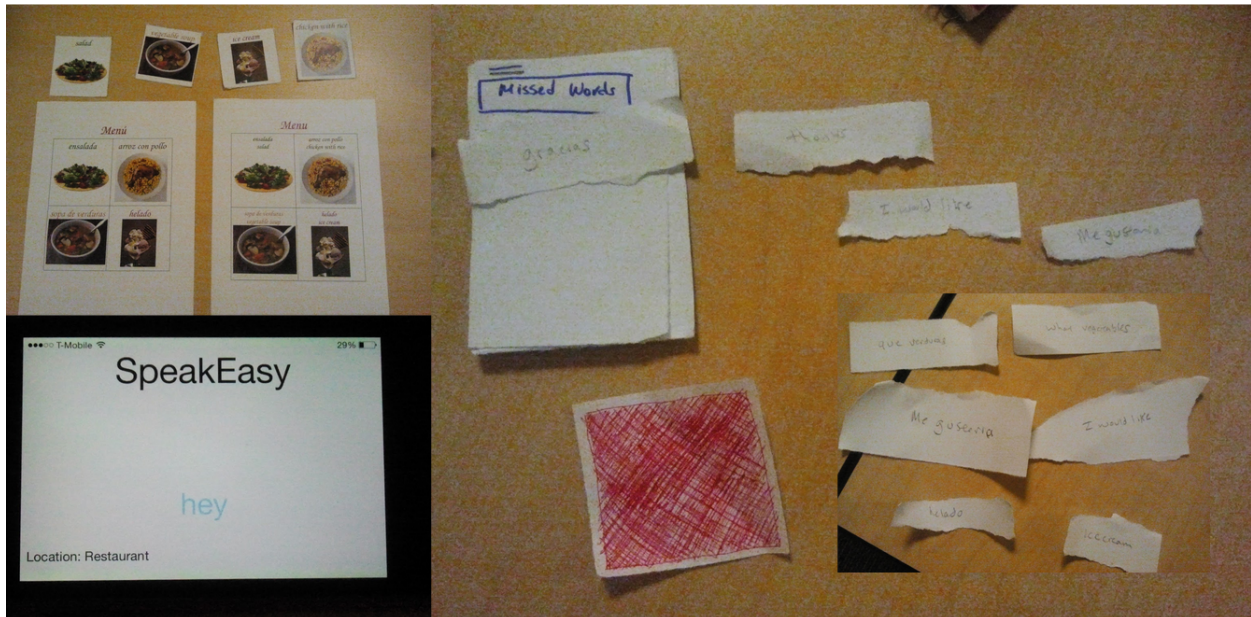


The mobile application was used in pairing with glasses and smart eyewear gestures to create a wizard-of-oz simulation that let users experience the context sensitive suggestions and intuitive pronunciation functionality of our application. The mobile application was displayed to the periphery of their vision and was used to display real time suggestions and translations based on the users current conversation. The user could make a forward swipe gesture along the rim of the glasses to bring up another suggestion. Simply saying the word "translate" brought up a translation of the currently selected phrase on the mobile application, and tapping the rim of the glasses played an audio recording of how to say the chosen phrase.



The glasses and paper cut outs were also used to allow for translation of objects in the users direct context. Users could use this functionality simply by looking at words in the foreign language with the glasses on, after which english translations would appear beside them.





METHOD

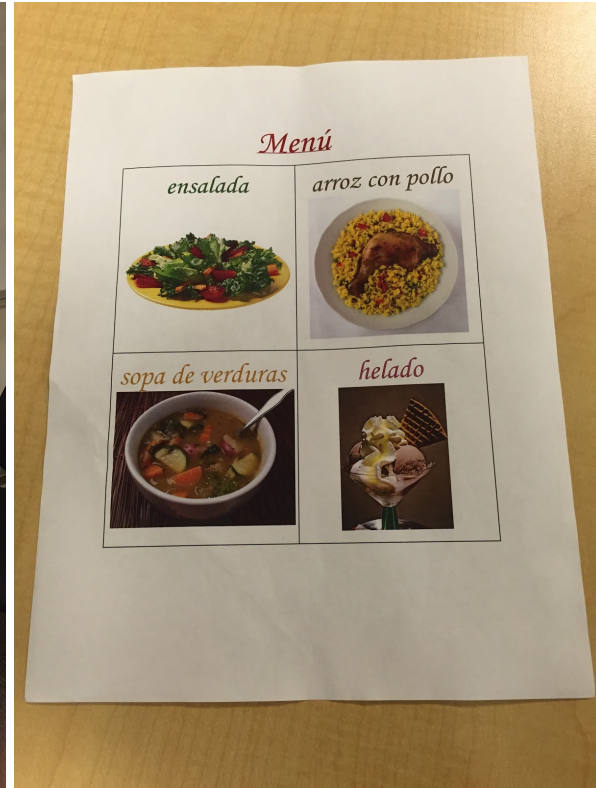
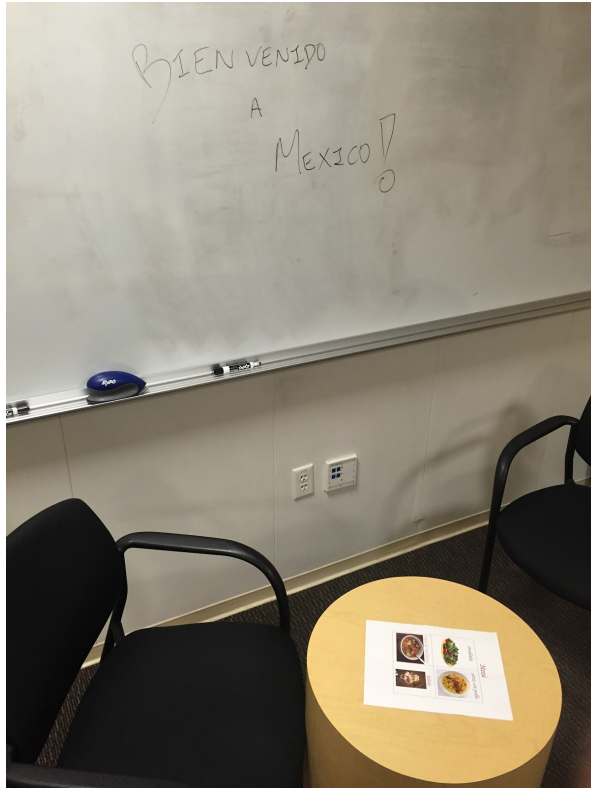
We designed a low-fi prototype to include both a paper prototype and a wizard-of-oz technique. We created a fictional restaurant scenario to simulate a situation in which our product could help our test user. We expected to get richer feedback about our user interface design like what worked and what didn't.

1. Participants

We recruited Stanford students that had studied some Spanish before. We tried to look for students that understood the basics of the language, but weren't very comfortable with it. We were fortunate to have participants that were willing to help us without being rewarded.

2. Environment

We set up an environment to resemble a restaurant in Mexico. We arranged a chair, table, and a menu in Spanish with some food options. To help set up the environment, Eric would greet our participants and explain the setting for the role play. Carlos also played a Mexican waiter that would only speak Spanish and whenever the participant mispronounced words would act as if he didn't understand.



3. Tasks

The three tasks that we tested were: intuitive pronunciation, contextually sensitive language support and active language learning integration. For each one of them we had scenarios prepared to induce the participant to perform those tasks.

3.1. Intuitive pronunciation

Whenever the participant had problem reading or pronouncing the word, he could tap on the temple of the glasses and the "glasses" would repeat back to the user the word in Spanish. The waiter would induce this task by verbally saying that he didn't understand what was said and expressing doubt whenever the participant mispronounced the word in Spanish.

3.2. Contextually sensitive support

Whenever the participant struggled with what to say, the "glasses" would show suggestions of phrases that the participant could say to the waiter. The participant could change options by swiping the temples of the "glasses", and see translations of

the phrases by saying “translate.” The setting induced the participant to ask questions about the menu and respond to the waiter’s questions about his or her order.

3.3. Active Language Learning Integration

After the conversation with the waiter, we would give to the participant an "app" with the words that he mispronounced during the interaction. On the "app", he could see what those words meant in English and check the correct pronunciation of the word in Spanish.

4. Procedure

For the experiment, we assigned specific roles for each member of the team. Carlos led the role play as the waiter in the Mexican restaurant. Eric was responsible for greeting the participants and kept track of mispronounced words as the wizard-of-oz behind the paper "app". Gabriela was the wizard-of-oz behind the "glasses". Tommy was responsible for recording the participants feedback and reactions during the interaction.

First, Eric would greet the participant. He would explain the project and how our UI worked. Then the consent form was given to the participant to sign. Eric would also set the stage for the role playing telling the participants, "you are in Mexico and you just walked in a restaurant."

Carlos took over and welcomed the participant to the restaurant. He would try to prompt questions and interact with the participant as if he was a Mexican waiter. He would only speak in Spanish, and he would try to induce the user to perform the tasks using our UI.

While the participant was interacting with Carlos, Eric would take note of words that the participant mispronounced. Gabriela would type suggested phrases on the "glasses" that were appropriate for the conversational context. The user could choose to use one of the suggested phrases, swipe to a different suggestion, or request pronunciation support. If the participant swiped on the glasses, Gabriela would type a different suggestion. If the participant tapped on the glasses, Gabriela would speak the word in English or Spanish to give translation and pronunciation support to the participant.

After the roleplaying finished, Eric would give the "app" to the participant so that he could interact with the app to receive feedback about his or her pronunciation. The feedback included words that he or she missed, what the word meant, and how to pronounce them correctly. We would also ask the participant questions about their experience with the UI, what they liked and didn't like about it, and what they would add or change.

5. Test measures

To measure the results of our test, we looked at how relevant a user found a suggestion, how many times each feature was used, the number of gesture swipes during the conversation, the fluency of conversation, and ability to pronounce or learn to pronounce words.

RESULTS

Our test users were all able to order food in our restaurant scenario, use phrases provided by the app, and review their mistakes after ordering using the paper prototype.

The main issue in terms of interface usage was that the glass gestures were difficult to remember. During the restaurant scenario, especially when confronted with the additional pressure of having to speak to the waiter, the testers all exhibited hesitation before using the gestures to navigate through suggested phrases. When testers did use the gestures, many used them incorrectly, mixing up the tapping gesture with the swiping gesture, and assuming that the "translate" command would translate the waiter rather than the phrase.

In addition, a couple of our test users felt that it was slow translating from the suggested phrase (in Spanish) to the corresponding translation (in English). This resulted in the conversation with the waiter being a little stilted and choppy. This was more evident in the users that new very little Spanish.

Users also felt that it was difficult at times to understand what the waiter was saying without translation. Thus the phrases suggested to them were not as helpful as they could have been, since they could not use the suggestions to respond appropriately if they didn't know what they were responding to.

A few users also felt that while the mobile app was useful in reviewing their mispronounced words, it did not give them as much information as they would have liked. Many were interested in how they had pronounced the word themselves, and where they were using it in the conversation.

There were several features the test users liked. The tests showed that the mobile part of the application was straightforward; after we demonstrated the basic gestures for scrolling through and reviewing the words they missed, they appeared to have no trouble with using it. One user also thought that the UI for the eyewear application was clean and uncluttered, which he thought worked well.

Overall, all of our users thought that our app idea was a good concept. They felt that it would be very helpful for language learners, and they could see themselves using it.

DISCUSSION

From our experiment and results, it was clear that while our application was found to be useful, some of the design decisions we had made were unintuitive for users. With the hesitant response from users and misinterpretation of the smart eyewear gestures, we are limiting the gesture interface to only a single gesture that allows users to swipe through phrase suggestions. The translation gesture, we decided, would be removed and instead the translation would be displayed together with the phrase. This would allow users to quickly and easily understand what the suggested phrase and remove the latency of translating from Spanish to English that many of our testers noted as a disruption in their conversation.

Users with very little Spanish experience also felt that it was difficult to understand the conversation. This made many of the suggested phrases ambiguous to them, as they didn't know what they were responding to. Given this, we want to add a subtitle feature to the next iteration of our smart eyewear display that will allow users to see real time translation of the conversation.

Users also expressed a desire to see the context in which they mispronounced words. We want to change the interface of our mobile application to allow users to review a transcript and recording of their conversations. This transcript would highlight the mispronounced words within the conversation. Users would also be able to hear the correct pronunciation, see a translation, and see a translation of the sentences preceding and following the word of interest. This will give the user both contextual information about the words with which they are having trouble.

APPENDIX

Script

Thank you for taking part in our test. Our app is called SpeakEasy, and it is a language learning app that makes use of both a Google Glass-like wearable interface as well as a mobile interface. The app will provide phrase suggestions and sentence completions in the language you are trying to speak, based off of your location.

We will first briefly show you the gestures for using our app.

(explain gestures for swiping through suggestions and translations)

Okay. For our testing scenario, we would like for you to pretend to be in a restaurant in Mexico. You will try to order food by speaking with the waiter in Spanish while using the app for suggestions. If you see this signal (*show red card*), it means that you pronounced a word wrong. After this happens, please use the app to correct your pronunciation.

(run through restaurant scenario)

Okay, now that you have successfully ordered your food, please use this mobile application to review and learn the correct pronunciations of the words that you mispronounced during the restaurant scenario. Here are the necessary commands:

(explain the gestures for going through a missed word in the mobile app)

That is the end of our test. Thank you so much for your time. We greatly appreciate it.

Critical Incident Log

<i>Incident Description</i>	<i>Participant #</i>	<i>Severity (0 low, 4 high)</i>
User looked at suggestions less frequently than others since they knew quite a lot of Spanish	1	1
User exhibited some confusion with using gestures to move through translations/suggestions	2	3
User had some difficulty asking what was in the soup shown in the menu	2	2
User moved through translations slowly because they were first presented in Spanish	π^*	3
User had some difficulty understanding what the waiter was saying	π	2
User expected the missed words to disappear after reviewing them	π	1
User wanted English translation of phrase first	4	3
User suggested that having a full transcript of conversation would be even better for review	4	2

**This user personally requested for his participant number to be π .*

Raw Experiment Notes

Participant #1

advanced speaker

at a certain point, it would be slow if the user is fluent enough

record grammar errors for more advanced speakers

interesting

important for suggestions to be on point

did not use suggestions frequently

Participant #2

Cool, Sick! - thinks the app is a good idea
rely heavily on translating what other person is saying
when translations go wrong, user might be frustrated (I didn't want to order soup!)
cool on the missed words feature, thinks is good idea
great experience
gestures not easy to remember at first
most confusing part was differentiating the gestures
translate what the other person says is most useful
phrases were useful
translate what the other person says is important
feedback was useful
save phrases that person didn't know and needed to be translated
grammar isn't necessary at first (maybe Spanish only?)
location and context is important (said so without us telling here explicitly that our app is geocentric)

Participant #π

gestures not easy to remember
hard for someone who knows little if suggestions come up in Spanish
conversation can become stilted if glass app is too slow
new users will have to rely on full translations much more
swipe to get rid of suggestions
listening is a challenge - can you understand what the other person says
need to have both translations at the same time, takes too long to swipe through
phrase and translation
(possibly have native translation speak and show learning language translation visually,
but must have it so that it doesn't interfere with conversation of person they are trying
to speak with)
possibly have user's conversation partner words translated
allow users to use the errors as a checklist instead
maybe suggestions aren't too helpful - for someone who doesn't know the language
too well, they might just say what is suggested instead of what they really want to say

Participant #4

Knows some Spanish
"Oh Wow", "Awesome" - the missed words app that shows users what they do
Experience was awesome

Seems like a good idea

Glancing at the app a couple times for help

user's native suggestions first would be more helpful

gestures are kind of unintuitive

suggestions feature was most useful

keeping track of progress - what other features?

it would be cool to go back and see conversation (like a transcript of how the conversation)

what other information would be useful?

keep the Glass interface basic, simple

would still be helpful to be able to translate what he says

Consent Form

The SpeakEasy 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 SpeakEasy. 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 (Gabriela Groth, Eric Chew, Tommy Truong, Carlos Araujo) 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 student researchers 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 SpeakEasy 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 _____