# TONGUES

Ishita Prasad, Alex Wu, and Anna Yelizarova

Tongues is an accurate, real-time translation application powered by the crowd. When automatic translators just can't get it right, Tongues enables users to ask the people around them what the *real* way to say a word or phrase is. This project is being developed by Alex Wu (Team Manager and User Testing), Ishita Prasad (Visual/Interaction Designer and Documentation Coordinator), and Anna Yelizarova (Developer and Visual/Interaction Designer). **This week, our mission was to learn how our users interact with our idea to build on it and suggest improvements.**

PROTOTYPE

Our prototype was composed of three sets of paper screens, one for each "task" that the user could perform with our application (the tasks are described in more detail in a later section of this paper). The user interacted with the screens with a pencil either "typing" by writing, or "clicking" by hitting a button with the pencil. When an interaction prompted a screen change or customized response, the human computer would swap out the screens or write on the screen, respectively.

The nature of our application lended a binary interface: one for the "asker" side, and one for the "translator" side:
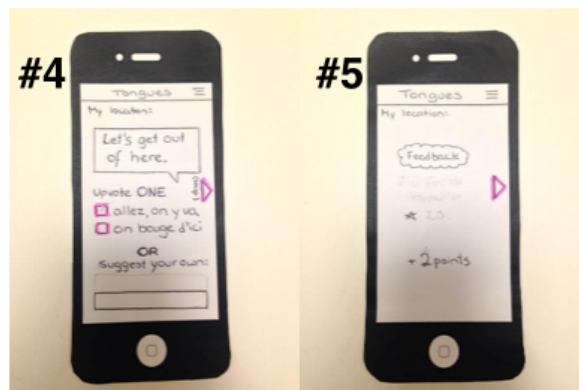
The first screen is the initial screen (#1), the second screen (#2) is what we displayed after 1 or more translations, and the third screen (#3) was a screen that allowed users to change location. In the first two screens, location was indicated at the top, and clicking the button next to it allowed users to move to the third "change location" screen. The other interaction element was a lone text box in which the user could "type" sentences.

On the translator side, we had an initial screen (#4) which showed a sentence to translate, translation options that users could upvote, and a text box where users could write their own translations. There was also a feedback screen (#5) which showed points the user had earned.

METHOD

Participants

Our participants had to be recruited carefully. The nature of our application required that each participant have some knowledge of at least one language besides English. Because a team member needed to act as a human computer during the experiment, we needed to ensure that the participants spoke a language that at least one of our team members knew as well.

Our first participant, Amanda, was recruited from a team member's class. Amanda had also studied abroad in France and learned French in an academic environment. She was also an iPhone user, and fairly familiar with the general interface of typical smartphones.

Our second participant was Kierstyn, a student who lived in the same residence as a team member. She responded to an email sent out to dorm lists asking for French speakers for a test study. Kierstyn is an advanced French speaker who has visited countries in Europe, including some that are French-speaking and some that aren't. She is also familiar with an iPhone.

Our third participant, Rebecca, was recruited from a language class (Italian). Rebecca has studied abroad in Italy and describes herself as being an intermediate/advanced Italian speaker, and is very passionate about language learning, too.

Environment

We chose a quiet, controlled, and consistent environment. The experiments were conducted in empty classrooms in a very private setting, sitting at a table. The participant and witness sat across the table from the ''computer''. Before starting the experiment, the team members tried to have casual conversations with the participant to put them at ease.

Tasks

As mentioned previously, our prototype experiment was designed around three major tasks that our application would enable, described below:

1) Get a colloquial translation for something you need to say
2) Translate phrases in other regions to learn a language
3) Translate interesting things in your language

These tasks ranged from easy to difficult, with Task 1 being the easiest and Task 3 being the hardest.

Procedure

During the experiment, one team member was the facilitator, who introduced the experiment, application, and dealt with other concerns; another team member was the human computer, who controlled the screens and responded to participants' interactions; the third team member observed.

The participant was first only shown the "asker" side of the app and were allowed to interact with it as long as they pleased (this averaged about three translations per person). In introducing the application, we told them to imagine being in the "target" location displayed (Paris, France or Florence, Italy), think of a word or expression in English that they weren't sure how to say in the target language, and to use our application to solve the problem (Task 1). We also asked them to further imagine needing a translation for a different language, and asked them to use our application to solve that problem, as well (Task 2).

Once the participants were ready to move on, we presented the "translator" side of the prototype, where they could upvote or write their own translations. Once they had done this, we would present the feedback screen which showed them how many points their answer chose. They would then repeat this process for different translations as many times as they liked.

Test Measures

The observer noted how the user interacted with the app. In particular, we looked out for clues about their comfort with the app, their confusion or annoyance, and their level of enjoyment or delight. In order to this, the observer noted the speed with which the participants interacted with various components, their facial expressions, body language, verbal comments, and other signs of confusion, comfort, and enjoyment. We also noted down any questions or comments that the participant had. Finally, at the end of the experiment, we listened to the participants' input about their experience with the application, both in terms of concept and task interface.

RESULTS

Participant 1

Amanda handled the app smoothly. She paused a little bit longer on some screens, and did not attempt to change the language on her own. When shown the translator screen, she asked "What is this for?" referring to the points, indicating that she did not understand. After we explained the points to her, she asked whether she would "level up" or if something different would happen if she earned enough points. However, in terms of buttons and interactions, she was able to smoothly interact with each one, and did not seem surprised or uncomfortable by any of the changes. Additionally, in her comments, Amanda told the team that she enjoyed the concept and was excited to see such a product. However, she also noted that she would not feel incentivized to continue using the "translator" side, and still did not understand the purpose of the point system.

Participant 2

Immediately after beginning the experiment, Kierstyn said that she wished she had such an app while travelling. She used the app very simply and quickly without asking any questions. She tried to switch her location from Paris, France to Montreal, Canada (both French-speaking) on her own, before we described Task 2. She laughed at the colloquial difference in the resulting

translation, indicating an element of delight and enjoyment. She performed several more translations, and commented that the the paper prototype was aesthetically pleasing. When we moved onto the translator side, Amanda easily and comfortably moved from one screen to the next, but vocalized a desire to win something.

Participant 3

At the start of the experiment, Rebecca noted how conversational Italian in Italy was different from her classroom learning. She seemed to enjoy the asker side of the app, doing more translations than required, and performed each asker task smoothly and without questions. Her interactions seemed to indicate that she understood the navigation and flow intuitively. She also vocalized her enjoyment of the app, and said that she could imagine the app being real, and said she would use the "translator" side of the app to help her practice outside of class. She also said that she would appreciate being able to compare herself to her friends.

DISCUSSION

There were a couple of (somewhat) major obstacles our users experienced while testing the product. The first - and arguably the most major - obstacle was related to our translation reward system. Our users could not understand how our point based reward system worked, and could not see themselves being motivated by it. We have come up with two main solutions to this problem: stronger gamification, and material benefits to translation. On the gamification side, we were thinking that we could switch to more of a badge system (rather than a point system), which would work similarly but would be a more tangible reward then points with no reference. We were also thinking that we could allow users to compare their badges/progress with their friends, which would enhance the value of the badges (people would value them more if they were competing with their friends for them). In terms of the physical benefits, we were thinking that a physical translation incentive might be more motivating than gamification is. This would involve partnering with airlines, cities, or other travel/language related groups to provide some sort of incentive system (some form of discounted or free travel) for the top users of the app.

The second major obstacle our users experienced was related to our UI - mainly our buttons and transitions. We noticed (and users commented that) our button system was slightly difficult to use, and that our transitions between tasks could be quicker (for example, after translating a phrase, you have to go through a couple of different screens to get to your next translation). As far as the buttons go, our solution is simply to design the buttons in a more appealing/intuitive way. This will allow users to use the buttons without thinking about it, and with minimal effort (a single step process rather than selecting an option and then having to click another button to continue).For the transition speed, our solution is to combine any extraneous screens and have immediate feedback on the main task screen. This will mean that users will be able to transition/repeat tasks immediately, rather than having to go through multiple screens to do so.

Using paper models limited our ability to see how users would use our app while traveling/on the go, which is one of our main use cases. However, this testing did allow us to find several flaws in the design of our product, and will allow us to continue to improve the product.

APPENDIX

Demo Script:

Facilitator: Hello, and first of all, thank you for participating in our user study. We're testing a prototype of our application, which you can use to get translations for sentences, and also translate other people's sentences. This [points to paper prototypes & human computer] is our prototype. We will ask you to use our application to solve 3 different problems, after we show you a brief demo of how to use the prototype. Do you have any questions?

Participant: No.

Facilitator: Awesome! Now, you can use this pencil to type [uses pencil to type in username on demo screen] and to push buttons [presses enter button]. Our human computer will make the screens respond appropriately to your actions [switches screen to "logged in" demo screen]. Do you have any questions about this process?

Participant: No.

Facilitator: Okay. So now, onto our first task: for this, imagine that you're in [Paris/France] and there's a phrase you want to say but don't know how. Please try to use our application to solve this problem.

[Participant does 1st task]

Facilitator: Onto our second task: now, you want to get a translation for something other than [French/Italian]. How would you do this?

[Participant does 2nd task]

Facilitator: Cool! Now we will switch modes, and go into "translator" mode. Use the app to translate sentences for other people.

[Participant does 3rd task]

Consent Form:

The Tongues 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 Tongues. 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 (Anna Yelizarova, Alex Wu, Ishita Prasad) 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 Tongues 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_____

Application Screenshots: