# One World – …Forget the Language Barrier

## Team Members

* Alex Cheng
* Angapparaj Kalimuthu
* Daniel Mejía V.
* Kenny Risk

## Overview

One World is intended to be a communication tool that enables people to take part in conversations with each other while each participant speaks in their native tongue. The goal of One World is to make possible conversations that were previously difficult due to language barriers, while exposing an interface that is natural and efficient. As we have discussed in our previous report detailing our inquiry, analysis, and design sketches, our focus is to design an interface that is viable as a communication device given the current imperfections of machine translation and speech recognition technologies.

## Prototype

Our project is focused on making communication possible and natural for people who do not share the same language. Given the open-ended nature of conversations and the need to see how natural people would feel when communicating through the system, we decided that a paper prototype would not be sufficient for the insights we were after. With a paper prototype we either needed a fairly scripted scenario that would not allow users to speak freely, or we had to introduce long delays for writing on paper the detected text, looking up translations, and writing back what the other party responded. Moreover, we would not have had insights on the kinds of errors that could be expected with current speech recognition and translation technologies.

For these reasons, we decided to implement a higher-fidelity prototype that would enable our participants to use our system more freely, naturally, and efficiently. The prototype consists of a Windows application that presents the interface we chose as part of our contextual inquiry analysis, and includes the functionalities we chose to support. Note that we deliberately ignored other interface sections such as the user profiles and the calling screens because our focus was on the actual communication and the features to support effective recognition, translation and error handling.

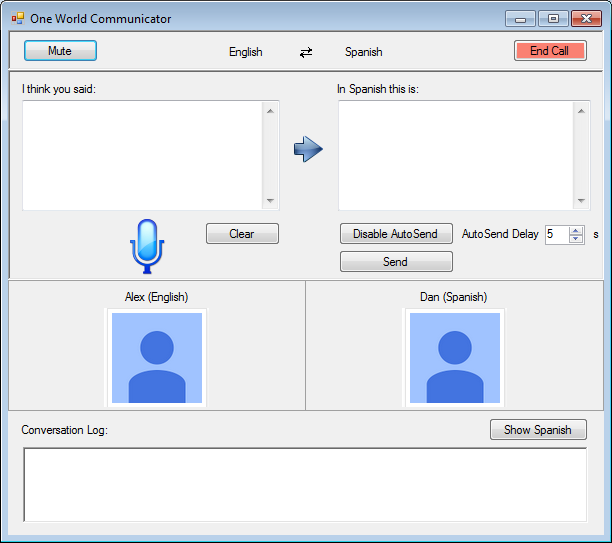


Figure 1: Main UI

As shown in figure 1, the interface includes the following sections and functionalities:

* Detected speech
* Translated text
* Auto-send messages (and ability control delay or disable)
* Manual submissions
* Clear option (for rephrasing totally wrong speech detections).
* Conversation log which includes sent messages and responses from the other party

Even though we wanted the prototype to be closer to the experience we envisioned, we knew that in order to test the specific scenarios and conditions we considered relevant for the design choices we needed to have full control over the interface and not depend solely on speech and translation technologies which can be highly inconsistent and unpredictable. Therefore, we created a mechanism for fully controlling everything that the participant would see when using the interface. Specifically, this functionality was added through a second screen which allows an “operator” to type what the user sees (both for the participant's speech and for the translated responses from the other party) which enabled us to introduce specific errors and scenario conditions consistently.

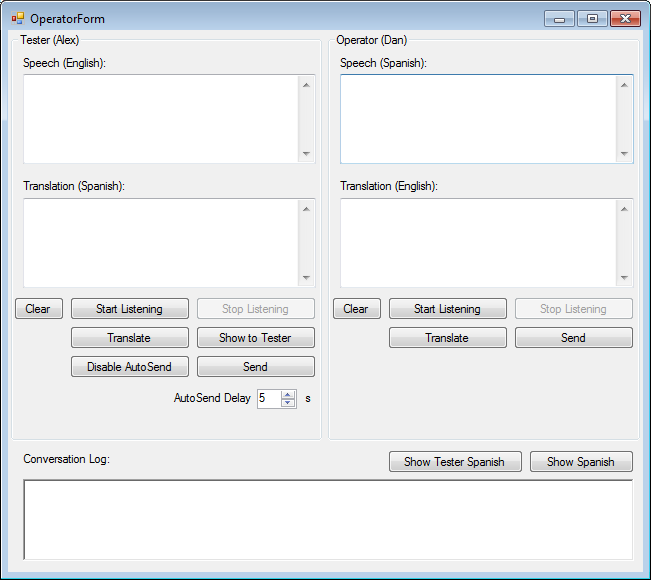


Figure 2: Operator UI

As shown in figure 2, the operator screen enables the following testing functionalities:

* Running speech detection (when the participant or the other party speaks)
* Translating text
* Manually typing the detected text (and introduce specific errors if needed)
* Deciding when to show to tester the detected and translated text

Additionally, for ease of testing, the “operator” screen provides the option to select what the user clicks on or changes in the main interface (e.g for submitting messages manually, or changing the auto-send delay).

Putting all of these options together resulted in a prototype that is interactive and provided us with the option of trying real technologies or efficiently faking them while testing the interface with real participants. Note that with this prototype, for practical purposes, the “operator” also plays the role of the other party in the communication.

## Testing Method

Having prepared our functional prototype, which as described above was specifically designed with the user testing sessions in mind, we proceeded to conduct our user study in search of usage data that could that we could use to drive revisions and improvements to the overall interface.

### Participants

We carried out our tests on November 16th, 2013 at Odegaard Library on the main campus of the University of Washington. We were able to convince three volunteers around the vicinity of the library to take part using Starbucks gift cards as incentives. The testers were all young, most likely undergraduate, students. Most people around Odegaard were trying to study, so naturally it took a while before we were able to find these three volunteers that were willing to help. For confidentiality purposes, we will call them Jenny, Katie, and Louis, respectively in the order of the test sessions that were conducted. All three were very nice (the kind of people willing to help with a study) but they had some differences as well. Jenny was extremely nice, and seemed to genuinely enjoy the experience. She was smiling throughout the entire testing session, and was even apologetic to the conversation partner when role playing in the face of difficulties introduced through errors in machine translation and by us. Katie was a little shy, and seemed a bit daunted at first, but adjusted and communicated efficiently and was candid in her feedback. Louis ran into more difficulties due to the language chosen for the study. While he wasn't apologetic like Jenny was, he was sincere in his approach to resolve misunderstandings while acting out the tasks, and was keen to offer his opinions about how he would improve the interface.

### Environment

We rented a quiet study room on the third floor of Odegaard and used that as the location where we would hold the testing sessions. As far as materials for the study, we brought with us a laptop, microphone, mouse, and external monitor and got everything set up prior to looking for participants. The room was much less spacious than what was hoped for and a bit narrow, but we were able to find a configuration that was sufficient enough for the tester and our four group members. After the second user test, our time at the room was up, so we used another similar room that was available for the third and final test.

### Tasks

As part of the contextual inquiries, we designed three elaborated tasks that illustrated important audiences and use-case situations in which we considered our proposed system to be very useful. The first task consisted of a business meeting in which a person is going over the terms of an important contract with the VP of a Taiwanese company. The second task consisted of making travel arrangements for world cup in Brazil. The third task consisted of giving a speech to a Japanese girl with whom you have been talking online for some time and would like to get serious with.

Even though these are our main tasks and remain illustrative of the kind of scenarios we would like to improve or make possible with our system, they were not very practical for user testing. The first reason is that they would take very long to complete and we would need the participants to spend too much time with us. This was important because, as we mentioned, our participants were volunteers found near the library. The second reason is that we wanted our participants to feel and act natural when using the prototype and asking them to act as part of a business meeting or online dating setting would not have helped. Finally, in order to conduct the user testing smoothly we needed to use a language one of us was familiar with and for which translation technology is currently good enough.

For these reasons, we decided to focus on a modified version of the travel arrangement use-case in which Spanish or Korean are used instead, given that Daniel and Kenny are familiar with them respectively. More specifically, the situation consists of a person planning a trip to Spain (or Korea) and talking to a travel agent to make the arrangements. For the purpose of testing we split the use-case into three main tasks that could be performed naturally by participants.

The first task consisted of *simple greetings with the travel agent*. The second task required more interactions and consisted of *booking a hotel*. The third task was more open-ended and prone to errors and it consisted of *finding about things to do and buying tickets for a soccer game if available*.

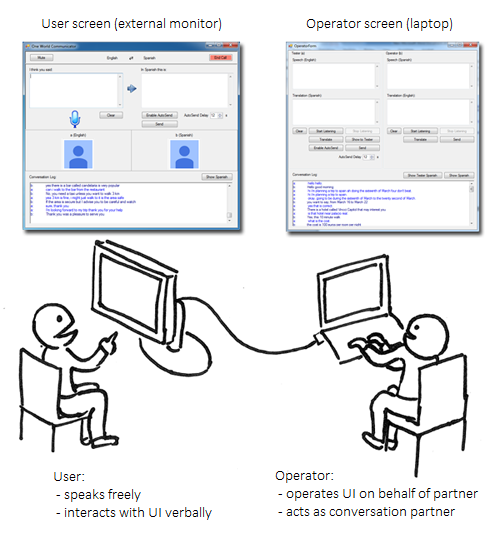
With these three tasks we aimed at making it easier for participants but also cover most of the spectrum of system functionalities including error handling. In addition to these main tasks, we added a fourth task to the studies for confirming the booking with a credit card which would bring closure allow us to test the use-case of numeric inputs. The specific tasks and notes are included as Appendix A.

### Procedure

In order to a have a predictable process and comparable results, we followed the same procedures with each of the participants and we tried to keep the interactions with the prototype as natural as possible.

The first step was to talk to the participant, and explain our project and what we were trying to accomplish. The second step was to hand them a sheet of paper which explained the use-case and the list of tasks to accomplish. The third step was to show them the prototype and explain the mechanics of it while encouraging them to talk to us about anything that felt natural or unintuitive. Additionally we would ask for permission for video-recording.

The actual mechanics of the interaction consisted of a laptop computer with the Windows application mentioned earlier. An external monitor connected to the laptop was used to show the main interface to the participant. However, there was no mouse or keyboard available to the user as the operator needed to use them, so the user would instead need to verbally tell us the actions to perform, similar to the way in which users interact with a paper prototype. Meanwhile, via the laptop, the operator would control the interface including everything the user would see as detected and translated speech, as well as performing the actions requested by the participants. In addition to this, the operator would also play the role of the other party in the communication (i.e. the travel agent). To simulate the video feed of the partner, we had the operator speak out loud what they needed to say before presenting the translated text of what was said to the user. We explained to the user that they could imagine the portrait icons as showing video feeds if this were a real product. Separately, two observers would be taking notes about the interactions and record video. Below is an image showing the interaction setup.



*Figure 3: Testing setup*

In order to make the communication predictable and consistent we decided to not use the speech detection option of the prototype during user testing. Also, there would have been practical difficulties in using it such as the need to explicitly (know when to) start and stop the speech detection.

So instead, the operator would listen to what the participant said, type it, translate it, show it to the user, and click send if the user requested so. Then the operator (playing as a travel agent) would type the response in the foreign language, translate it, then say it verbally for the participant to hear it and finally show the message text to him or her so that it felt like a quick translation of what had just been said. Actually having machine translation available to us in this manner proved helpful both in terms of making the interactions quicker, and introducing realistic errors. Finally, the operator was also in charge of introducing errors. Since he was in control, he would decide to introduce a speech detection error (typing something different), a translation error (changing the translated text, and as a travel agent asking for a rephrase), or a speech/translation error on the other side (sending a translated response that would not make sense to the participant).

The last step was to talk to the participants about how they felt and ask them for feedback. For the first two participants Daniel played the role of the operator/travel-agent, Kenny was the facilitator and Alex and Raj were the observers. For the third participant, Kenny was the operator/travel-agent, Alex was the facilitator and Raj and Daniel were the observers.

### Test Measures

Each test took approximately thirty minutes of time. The main thing were looking for with the studies was to understand how different people would communicate with our application. Overall, we wanted to see if participants felt natural and were able maintain fluid and effective conversations. More specifically, we wanted to find if they would find the interface intuitive and easy to use. We wanted to observe how they would react when there were speech detection errors and how they would work to solve them. Would participants find it natural to make small editions directly on the text? Would they rephrase the whole sentence? Moreover, we were also looking at their reaction to responses from the travel agent that did not appear to make sense. Lastly, we were also interested in details like the discoverability of the manual send, and the clear options, and more importantly, we wanted to get an idea of how useful the auto-send feature really was and whether people would prefer it as the default behavior. At the end, we asked participants questions about general impressions and issues that came up during testing because we wanted to gather general feedback directly from their experiences. All of the collected notes as well as the conversation logs are included as Appendix B and C, respectively.

## Testing Results

One of our goals was to observe how different personalities would react to and handle the different error conditions that are common with speech and translation technologies. Before engaging real participants, however, we wanted to use the prototype with real speech detection and find the kind of errors that are possible or even common when speaking naturally. We found many problems that can easily arise when there is ambiguity or the overall context of the full sentence is not understood. For instance, we found the Spanish word "sí" (which means "*yes"*) to be detected as "si" (without the accent) and whose translation is "*if".* During the conversation, this lead to Daniel responding "*yes"* but Raj seeing "*if"* instead and therefore asking "*is that a yes?*" and Daniel responding in Spanish "*Yes. It's a yes"* with the translation sent as "*if an if"* which was funny and made no sense. A similar situation arose with the word "*location*" which was translated to "situación" which sometimes might mean "location" but more commonly means "*situation*" and therefore it lead to the sentence "*what is the situation of the ticket*" Additionally, we found that punctuation was not detected correctly and this lead to some loss of meaning and sometimes was the cause of some of the translation errors.

We then proceeded to having real participants go through the different tasks and we found that they had different personalities and reacted differently to certain situations and errors.

Jenny, the first participant, was extremely nice. She would smile all the time despite the errors that arose in the process. She even apologized for some of them. Initially she had a hard time catching errors and she would allow the incorrect messages to be sent, but after the responses made it clear that there were problems with her messages she started paying more attention and corrected the messages before sending them. We were pleased with this and consider it a natural way to find the need to correct errors. The other problem we found was that sometimes it was not easy for her to express when there was something that did not make sense. For instance when the travel agent meant to say "your card was accepted," he (as operator) voluntarily introduced an error and sent the translation as "your bag was accepted." Jenny naturally repeated bag in English as a question, and if this were a real conversation the travel agent would be confused at why the person is asking about bags. This finding was important because it hinted at having a better way of signaling errors which we had already started to think about. Finally, Jenny said that she preferred sending the messages manually (no auto-send) but she mentioned that younger generations would appreciate it (even though she was very young).

Katie, our second participant, started with a very long speech but then naturally started using shorter sentences which we were expecting and hoping. Katie was more precise and she became very efficient at spotting errors and correcting them quickly before sending. At first, she did not notice the "send" button so she relied on auto-send, but when she noticed it she began using it every time. She also noticed the clear option and used it for completely wrong detections. Katie did not seem to enjoy the experience as much as Jenny, however. As part of her feedback, she mentioned that it did not feel like a conversation. While this was expected to a degree, it also points out an inherent limitation of the interaction model where the user must be afforded opportunities to correct errors. One the other hand, Katie mentioned that she valued the ability to make corrections before sending. With that said, when asked whether it would feel more conversational if speech detection results were returned more quickly than the time it took for us to type out her speech, she said that it would. This also validated our decision to use a higher-fidelity prototype because a paper version would have been much more unnatural. As a final comment, Katie said she did not see the point of being able to specify the amount of auto-send delay and after we explained she suggested having only an enable/disable option.

With Louis, our third participant, we used Korean instead of Spanish and this lead to more translation errors that were difficult to deal with. For example, the travel agent's introduction was misinterpreted by the machine translation as, "I called Kenny," as opposed to "I am called Kenny." After the testing session was over, he specifically pointed out such difficulties, but also mentioned he expected errors because of the translation involved and most of the time he could make sense of the responses even if they were not fully correct. Another thing that we did differently in this session was disable auto-send from the start (previous testers had disabled it for the most part, and there was a bug in the implementation). When we asked whether Louis noticed the auto-send option after the test was finished, he mentioned that he did but that he chose to leave it disabled as he felt comfortable pressing "Send" manually like in instant-messaging applications. As far as general impressions, despite the low quality of the translations, Louis stated that he appreciated the fact that such a tool would enable him to be able to talk to someone native to a foreign area (e.g a native travel agent knowledgeable with the area). Finally he said that to make sure that the translated text to send was correct, he would just copy it to a google translate window and verify it. This was important because he basically suggested at least an option to have the back translation feature we had originally discarded.

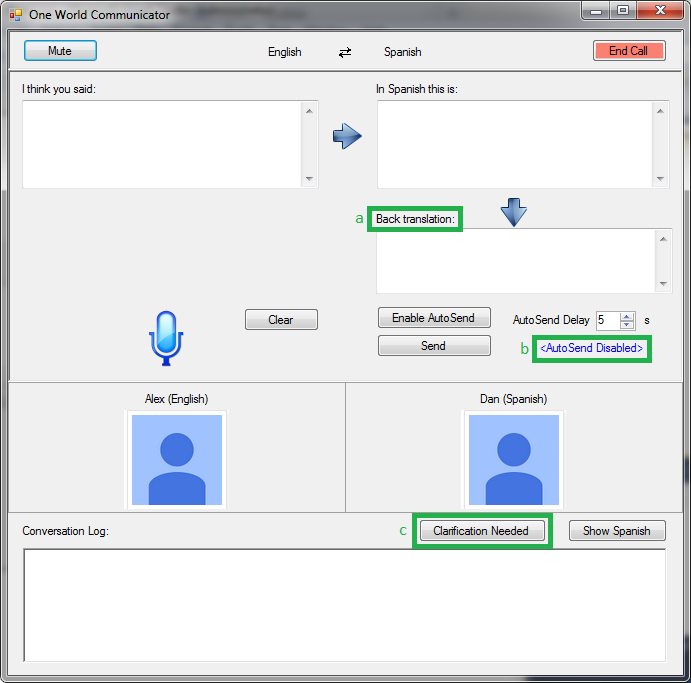
## Interface Revisions

One of our first findings during our testing session was that the auto-send feature was not as natural as we thought. Two of the participants did not use the feature until we pointed out it was available. One participant commented that she would prefer to keep it turned off since she did not trust the translation enough. Another commented that the timeout feature was not needed since she would either turn it on or off and did not foresee the need to change the auto send time interval. We do feel like more extensive testing would be needed to understand the usefulness of auto-send, but for now since user testing pointed us in that direction and since it can be confusing, we decided to disable it by default. The feature will possibly be handy for advanced users that already who have developed trust in the application.

The second finding was that sometimes when the meaning of a sentence was not clear, it was not easy to ask for clarifications because the cause could be a translation error not visible to the other party. When the original intent is lost, it becomes impossible for the recipient of a badly translated message to convey to the sender where the error is by using speech alone. This became apparent with issues like translation of “si” to “if” or with messages like "your bag was accepted" (instead of card) in which there is clearly a portion of the message that is wrong but asking about the translated version would not help. One solution is to allow users to flag certain ambiguous or incorrect words or phrases so that the other party is notified about something being unclear and the problematic section is visible in the original message. We talked about this in our original proposal but did not add it to the initial prototype. Based on these observations, we decided to include it in our final version.

Lastly, we saw that in some situations incorrect translations would cause the original intent to be lost and there was no way of verifying it before sending. This is especially important in languages where the quality of translations is lower (e.g Korean). One participant stated that he would copy paste the translated text to google translate and make sure that it was correct which is basically the back translation feature that we had in one of our alternate designs. This is a feature that we felt advanced users would appreciate, but we did not add it to selected design since we felt it could be too distracting. Based on current feedback, we are of the opinion that providing this as an option would greatly enhance the user confidence especially in less common languages.

Another thing to note is that participants tend to speak in longer sentences and sometimes continue to say something after briefly pausing. This introduces the need to show translated text as the user speaks, preferably when they pause. One could think of detecting granular sentences and enabling the user to send them together or separately. This would help them build up longer sentences and thoughts before sending them out, and correct errors only where needed. At any rate, the prototype in its current form does not support this ability to obtain detected speech in a more granular fashion. Thus, when a user speaks a very long thought, there is no feedback that the speech is being detected until the final text shows up. More importantly, having a long piece of text show up at once requires much effort on the part of the user to review before sending it. We thus think that granular speech detection should become a part of the final interface, but did not get a chance to include in our video prototype.



*Figure 5: Interface revisions. a) Back translation feature to detect translation errors b) Auto-send disabled by default c) Error signaling for requesting clarifications*

## Video Prototype

We had an interesting dilemma when trying to decide what kind of prototype would best illustrate our ideas. As part of our design process we had already built a functional interactive prototype, but the technical limitations of the prototype and the fact that it is was not a web application meant that it was not easily shareable. It also did not help that the learning curve to operate the prototype is a bit steep.

Due to these considerations we decided to do a video prototype that will showcase the interactive prototype that we build. This allows us to show a higher fidelity experience than a paper prototype while at the same time permitting control over the experience to a high degree, thereby letting us surmount the technical challenges with our prototype.

### Scenario and Tasks

We opted to use the travel arrangement scenario for the video prototype since this is a situation that we felt will resonate with a majority of our viewers. For the purposes of the prototype, we had Daniel play the part of a travel agent and Alex to play the part of the customer. Since Daniel speaks fluent Spanish, he was a natural fit for the travel agent role where he is communicating solely in Spanish. Alex is a regular traveler and is passionate about football, which let him play the role of the customer in a very convincing manner. We shot the prototype over two sessions held at the Microsoft City center building in Bellevue and the classroom at UW.

### Scene 1

We wanted to start the video by providing some context of the problem that we are attempting to solve. The first scene shows Alex trying to book a trip to Spain by calling an agent on the phone since he is fed up with planning trips on Expedia and values the ability to speak to a real person and get personal insights that are not possible online. Unfortunately, this ends in failure due to the language barrier between Alex, who speaks only English and the agent who speaks only Spanish.

### 

### Scene 2

As Alex continues his search, he discovers a travel agent who has a OneWorld ID. An excited Alex then proceeds to make a call to the agent through OneWorld.

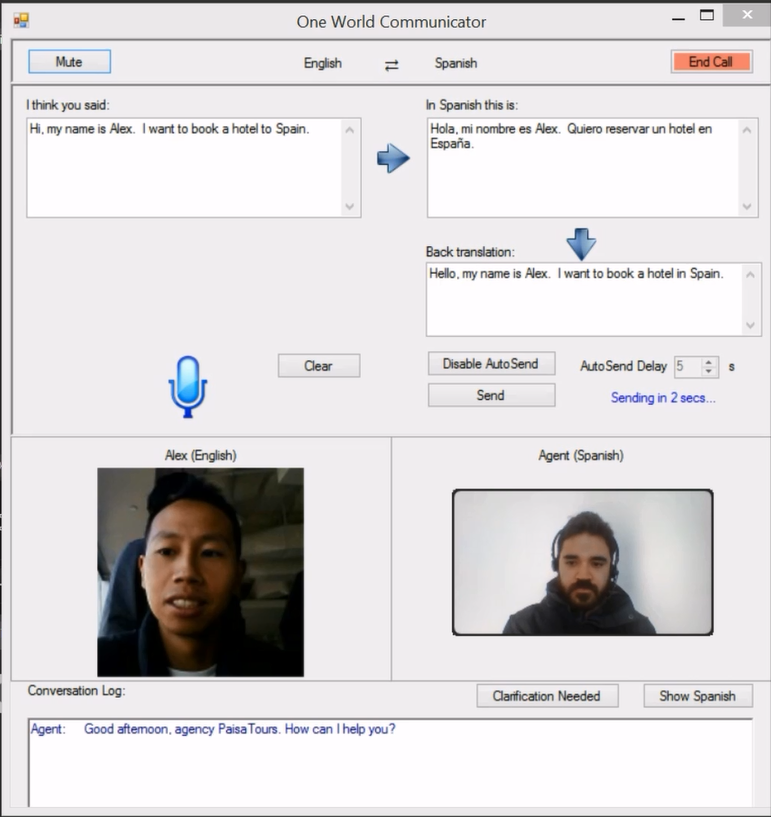


### Scene 3

This is the main section of the prototype, and it shows Alex and Daniel working through several tasks related to the travel booking scenario using the OneWorld interface.

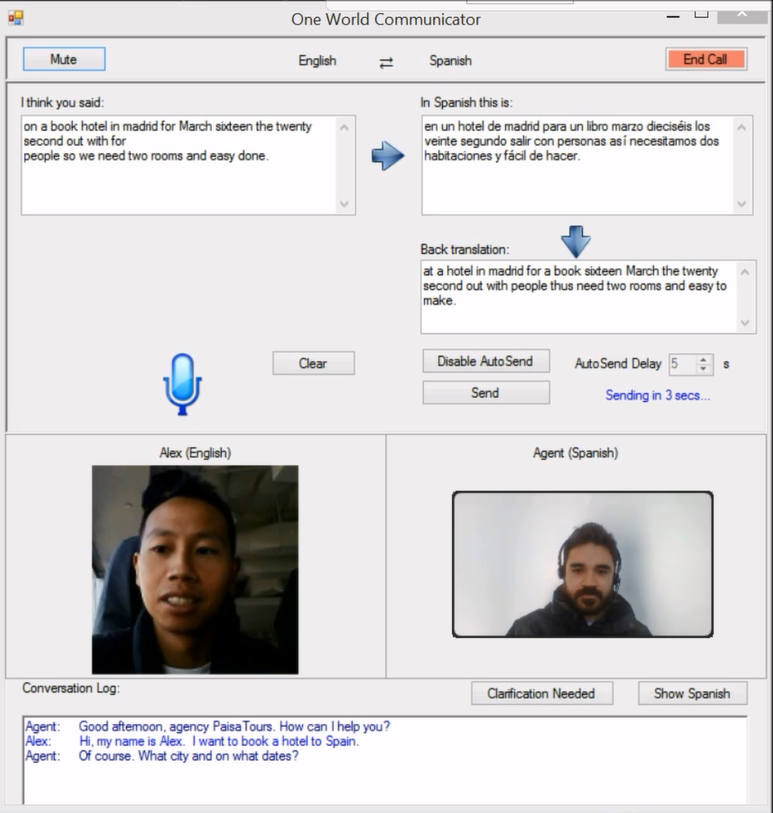
##### Task 1 – Basic introductions

Alex and Daniel greet each other and proceed with the booking process.



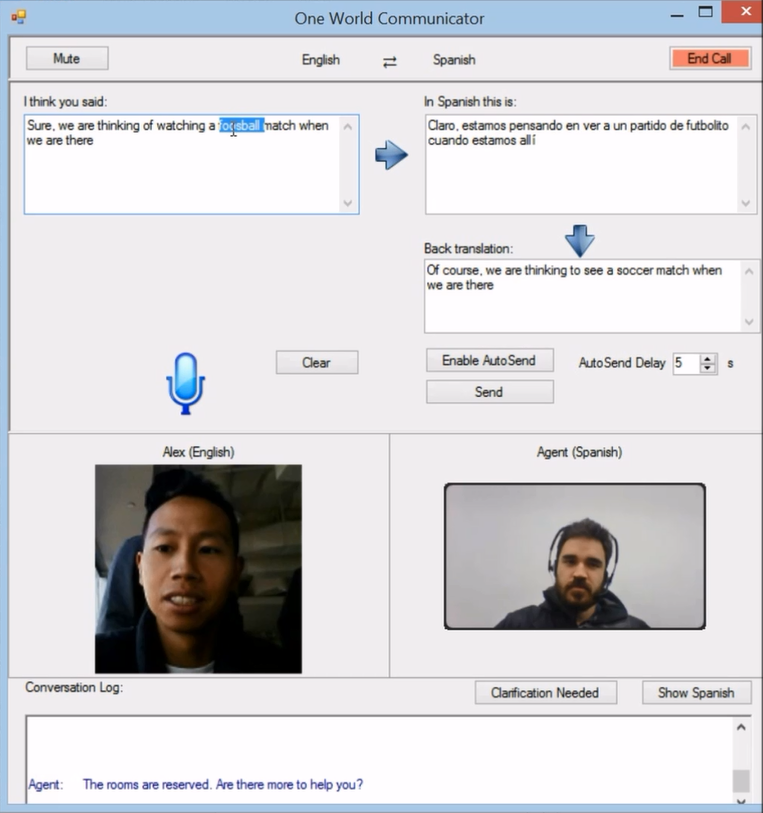
###### Task 2 – Hotel booking

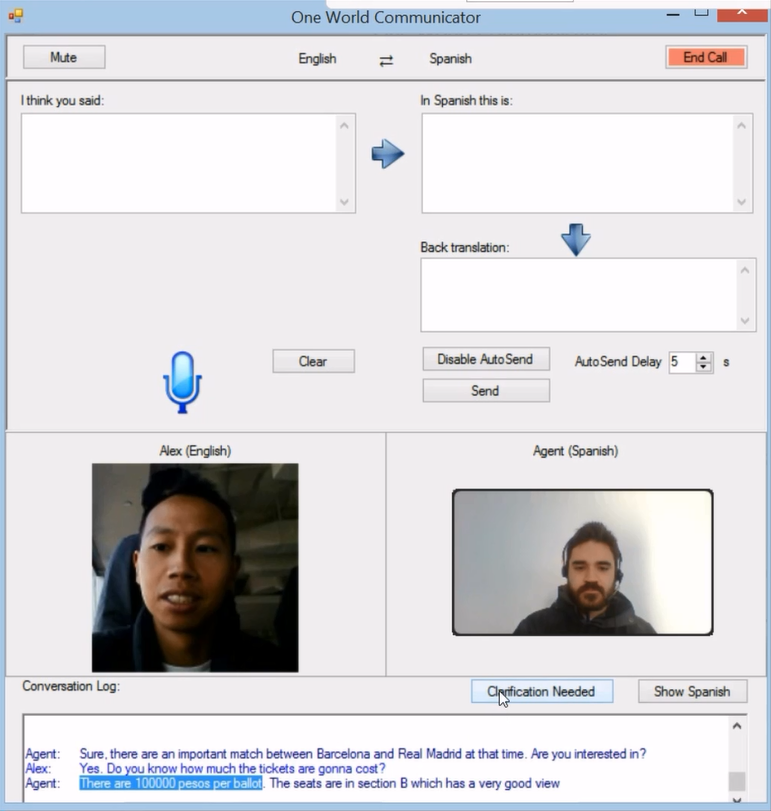
Daniel books a hotel for Alex. This task involves longer, more complicated exchanges as well as correcting errors introduced in the translation. Alex also uses the “Clear” option to remove incorrectly recognized text and rephrases his question in shorter sentences.



##### Task 3 – Booking tickets to a football game

This task is comparatively more difficult to complete and it highlights the corrections to speech recognition and translation on the interface (eg. Foosball to Football). We also show an example of using the “Clarification Needed” option which highlights questionable parts of the conversation to both parties so that it could be re-phrased or corrected.





### Scene 4

The last scene shows Alex enjoying his trip in Spain where he runs into Daniel at the football match. An amusing exchange then follows.



### Tools Used

The primary tool used for our video prototype is the high-fidelity windows prototype that we built for our user testing. A revised version of this prototype was created based on the findings from our user testing and it was used exclusively for scene 3. Since our prototype did not support video streaming we used Skype to supplement it. We setup a Skype call between two laptops, with Daniel on one machine and Alex on the other. We were running our OneWorld prototype on the machine Alex was using with the Skype video overlaid on top. This allowed us to simulate a real-time video call in an easy manner.

We used a mobile phone camera to shoot most of the external perspective video scenes. For scene 3 we used a screen capture tool called Microsoft Expression to record Alex’s interactions with our OneWorld prototype. These video clips were then encoded and mixed together using MovieMaker software.

## Discussion and lessons learned

Overall we are very satisfied with the observations, learnings, and interface revisions we did following standard human computer interaction processes, and we feel confident with the final outcome and design decisions. There are still open issues and questions to answer that would require more extensive and diverse user testing, but even though the open-ended nature of communication required us to scope the problem we believe to have achieved the goal of creating an interface to enable people who speak different languages to communicate effectively and deal with the limitations imposed by current speech and translation technologies.

The first important learning for us was that in order to create an effective solution to a problem, one must understand the current state, possibilities and limitations of the underlying technologies. In our case, we knew from the beginning that there were some limitations with current speech detection, and to a lesser extent with translation, but it was highly valuable to really understand the frequency and the kinds of errors that might arise and some of the causes for them. Learning and directly using speech detection and translations tools lead us to significantly focus our studies and our design at dealing with errors naturally and efficiently. Besides frequency and kind of errors there were a couple of other important findings. The first one was the difficulty with detecting punctuations and how this lead to translation errors, and the second one was the high variation of quality across different languages and how even though languages like Spanish might work well, Korean and many other languages require the interface to provide additional options to make the provide the user with confidence of sending the right information (e.g back translation).

Another important finding was that of not relying solely on intuition but also delay some of the decisions and let them be guided by data and observations from contextual inquiries and user studies. One important example of this was our original plan to create a speech-to-speech system that would output translated voice. Through contextual inquiries, it was easy to see that people value the ability to see the actual text and they also noted the importance of hearing the original tone of the voices even if the actual content was not clear until seeing textual translation. We later confirmed this design choices in our user studies.

In addition to this, we found the relevance of choosing the right abstractions needed for testing a particular design which has certain interaction requirements. In our particular case, even though we could have created a paper prototype, the unavoidable delays and the need to create a scripted scenario would have significantly hindered our testing because of the importance of fluidness when communicating and only a bit of extra effort provides us with a much more valuable interactive testing tool. We also understood the importance of designing and iterating on well-balanced tasks that are easy to explain and perform by participants but also provide sufficient coverage of the important situations and features in our interface. More specifically, we found the need to adjust and scope down our tasks according to the audience and this case make them less time-consuming to be able to convince random people to participate in the study.

Finally, even though we feel confident with our decisions and provide reasoning for them in the results and interface revision sections, we are aware of certain gaps and open questions that could not be solved with low-cost techniques for prototyping and user-testing. We particularly feel that we did not have enough coverage in terms of number of people with different personalities and types of situations they can find depending on the ways they communicate. We did gain important insight from seeing three different people interact with our prototype and react differently in similar situations, but we also know that studies can be much more thorough and accurate when considering more people from different age groups and backgrounds. For instance, even though our studies clearly pointed towards removing auto-send, there was some ambiguity and a higher-scale study could confirm this or point us in a different direction. We also noticed that the attitude and attention of some situations was affected by the awareness of this being a test (e.g us providing an extremely high price for the soccer game went unnoticed by one of the participants), but we accept this as a general limitation of this type of user testing.

# Appendix A: Task List and Operator Script

## Task List

*You are booking a trip to Spain over spring break with four of your friends. You plan to be in Madrid for 7 days from March 16, 2014 to March 22, 2014. You have to book a hotel and want to see a football (soccer) match. You will be speaking to a travel agent over the computer. What you have in front of you is a speech translator system. Simply speak to it in English.*

*We would like you to complete the four tasks below. Treat this as if you are booking a real trip, making sure you get all the information that you need. As you go through the tasks, feel free to ask the agent questions or clarifications. He understands that you are using a computer translator system which is not perfect.*

### Task 1

Greet your travel agent. Exchange names.

### Task 2

Book a hotel for you and your friends near Palacio Real. Make sure you gather all the information that you need.

### Task 3

You and your friends want to see a football (soccer) match. Get the information and book the tickets. Make sure you know how to get to the stadium.

### Task 4

Pay with your credit card and confirm the booking. Your credit card number is 4535 1245 5325 1804.

## Operator’s script / comments

Idea is to give them only information they ask. If they don’t say anything, ask if they have any questions. Hopefully that’ll be enough.

### Task 1:

* Name
* Who we are
* How can I help you?

### Task 2:

* Hotel info: Vincci Capitol
* Address: Calle Gran Via 41 28013 Madrid Madrid
* Phone #: 34-915218391
* 10 minute walk to Palacio Real; ask if it is close enough
* Each room has two single beds; ask how many rooms they need.
* Each room is $100 euro/night
* Internet, swimming pool

### Task 3:

* There is a match between Real Madrid and Barcelona on March 20th
* Tickets are at the upper level, $150/ticket
* Take the taxi to get there. Taxi will be about 30 euro. Ask the taxi to take Cuesta de San Vicente to avoid traffic.

### Task 4:

* This is to see how they handle numbers and finalize the trip planning.

## Korean Variant

For the third test in which Korean was used as opposed to Spanish, the same general scenario and tasks were used, except with different location names. In addition, instead of a soccer match, the user was to express interest in watching a K-Pop concert.

# Appendix B: Notes collected during User Study

11/16/2013 10 am - 2 pm  
UW main campus, Odegaard Library

## Dry test run with group members

Raj - tester

Daniel - operator (Spanish)

Auto-send: user thinks that you have to disable auto-send before being able to correct the txt. But the autosend becomes disable for the whole conversation.

Not clear how you can build longer sentences - the UI accumulates text without clearing the built up buffer.

translation error: "si" gets translated to "if" instead of yes (because of not having an accent), the agent tries to say it again, but it was still translated incorrectly, so user

We have decided to go without the speech recognition due to stability and operational difficulties (handing over mic, start/stop translation, shared mouse/keyboard between assistant (Kenny) and operator). It is easier for the operator to simply type out what is said, press translate and send to tester.

## Test Subject #1: Jenny

Language: Spanish

Daniel: Operator

Booking hotel and soccer match scenario

11:10am start

She speaks a little bit of Spanish

Long sentence for her introduction and request

Response from Dan

M: Is that near Vinci capitol? What's the price range? (two questions in one)

Resp

M: All right that sounds pretty good

(2 bed..2 queen … uh uh.. Street… are there any bedroom like that open?)

Long sentence… she left autosend to send out

Resp from Dan

Smooth so far.. Perhaps not so much speech recognition error introduced)

"Change teams to team please"

soccer

I'm so sorry… futbol match? Is there any futbol matches happening in the area?

"book them" - tickets

So far so good, no real difficulties

Stadium and I need four tickets

Didn't understand - response from Dan

"so that's four tickets, as in the number"

"Oops"

Slope of vicente - could you define, is that where the stadium is?

"Remaining Certa" - she understood this as near Certa

I'm so sorry I think I misspoke

Name and friend mistaken

She's very polite and nice. "I'm so sorry, I am having a hard time speaking today." (despite not having said anything unclear - the way the speech was detected was at fault)

Confirm booking

Speaking credit card number… very deliberately and slowly

I would like to change the number

15 minutes since start

"My bag was accepted?"

"Problem to trouble" .. Using the editing feature

Getting

Punctuation - voice recogntion

Translation

Exit interview:

- How did you feel about the experience? Really enjoyed it (she was smiling the whole time), and it was useful. "If you were to market it, it would find success."

- Jenny didn't find the autosend useful. She felt more comfortable clicking send. But noted that for younger generations (even though she looked quite young) they may choose to use autosend.

## Test Subject #2: Katie

Language: Spanish

Daniel: Operator

Booking hotel and soccer match scenario

11.47 AM Start

Katie..

C: Hi, I am Katie..

Continues to say " I will like to book a room, 5 people, near P.R"

Continues to say " we will be visiting --<dates here>"

Did not notice the send button at all.. (auto send kicked in)

She finished three sentences before Dan got a chance to speak

d: Dan informs about the hotel,,

c: How many rooms, single bed rooms (says wait) then says two single bedrooms

Notices translation error. Changes "brooms" to "rooms"

d: all are going to be single rooms, is it ok

c: Asks for total cost.. (notices translation error, "trust" to "cost"

Still did not click send button, always using auto send

d:dan gives the cost

c:she is not sure if it is for one room or for all six.

Notices translation error and asks to change part of the sentence

The wrong sentence was auto sent before correction could be made. :(

<user took some spanish classes, so thinks the incorrect sentence does make a little bit of sense)

Dan gave total cost and she says "that will be fine" repeats the sentence because of translation error. Still does not explicitly click "send". Always waits for auto send to kick in

Dan says "more than you want" and she responds with "is there a pool in the hotel" ????

Dans response has "If" but user still got the meaning..

She says "That’s it" Does not realize there are more tasks. Dan pointed it out to her.

C: Asks for football stadium directions. Notices incorrect recognition of "baseball" and ask for change to "football"

12:00 PM - 13 mins into the interview

Dans gives a really long response.

c: asks about the teams that are playing. Says "send" this time. Think she figured out the on demand send option.

User rephrases the query when Dan responds that he does not understand what she said.

User corrects some words and sends it.

Dan gives the option for barcelona match.

C: asks if they take mastercard. Dan gives bad translation, she clears and rephrases the query

D: asks for credit card number

User gives 4 numbers and waits for translation. The translation is mix of number and words. The user edit in place and changes the words to digits.  
User sends only the 4 digits after correction.

D: Says the bag was accepted. User did not notice the error. (maybe noticed and ignored it)?

D: Dan says thanks and finishes the conversation.

She feels the experience is not very conversational. Does not seem like speaking to a human

Would reducing the delay (of speech recognition) make a difference? - Yes, people will use it more

Is it better having option to click send? - Think it is a good idea to be able to correct it.

Some parts where you moved on despite the translation being less than perfect? - She understood the main point and decided to move on. Maybe because she knew some spanish already.

Anything that is not intuitive? -User asks why is there auto send delay? (Dan explained the reason).

User feels the ability to change the autosend delay is not really needed. She will either disable the auto send or use it.

User thinks rephrasing incorrect sentences is better for her. Maybe it would be better to type in longer sentences!!.

## Test Subject #3: Louis

12:34 Start

Language: Korean

Similar scenario as first two, only with soccer match replaced with K-pop concert, and different locations

Auto-send disabled this time due to buggy behavior.

Greetings. Kenny instructed that once Louiss think the recognition is good, just click / say Send.

Kenny introduced himself, but looks like the translation is bad.

Louiss try to clarify "No my name is Louis, what are you called?"

This time, it looks correct.

Long sentence to ask about booking the hotel for 4 ppl, asked about price and amenities.

Kenny answered "we have a hotel near Kangnam", but the translation turns out to be"Kangnam near hotel".

The translation for "how else can I help you" to "besides I help you". Louis decided that he understood and not need to ask for clarification.

Louis spoke multiple sentences all at once. Want to see concert, any artist and price.

The recognition was not good, so Louiss did 3 corrections (instead of clear / rephrase). After the correction, it was sent.

Louis confirmed the trip by asking about the cost.

Kenny's reply doesn't make sense. Kenny suggested (in English), that perhaps Louis could rephrase.

Louis rephrased by saying "No how much is the trip going to cost"

Louis corrected "book end" to "booking"

Saying the credit number, the recognition for the last 2 digits were wrong. He corrected it, by "typing" it out in words (five instead of 5). [since this is consistent with the rest of the correct speech recognition]

This test lasted longer than the first two, due to the difficulties that the machine translation presented.

Feedback:

* It was useful. I definitely appreciated the fact that I could get information from a travel agent native to the area as opposed to an American travel agent who would be less knowledgeable about the area
* Slightly frustrating but understand the translation
* Pressing "send" is OK, natural like IM
* When he sees the translation box, he said he would copy and paste that to google translate to make sure what he said is correct. (this seems to be a reference to back translation)
* He actually suggested back translation! We noted that we actually had another version of the prototype that has this in it.
* Notice the auto-send feature, but didn't want to use it because of error

# Appendix C: Conversation transcripts of tests

## Dry test run with group members

|  |  |  |
| --- | --- | --- |
| User | English | Spanish |
| a: | hello hello. | Hola Hola. |
| b: | Hello good morning | hola buenos dias |
| a: | hi i'm planning a trip to spain ah doing the sixteenth of March four don't beat. | Hola estoy planeando un viaje a España ah haciendo el 16 de marzo cuatro no le pego. |
| a: | hi i'm planning a trip to spain. | Hola estoy planeando un viaje a España. |
| a: | okay. going to be during the sixteenth of March to the twenty second of March. | bien. va a estar haciendo el 16 de marzo el segundo el 20 de marzo. |
| b: | you want to say, from March 16 to March 22 | quieres decir, del 16 de marzo al 22 de marzo |
| a: | yes that is correct. | Si eso es correcto. |
| b: | There is a hotel called Vincci Capitol that may interest you | Hay un hotel que se llama Vincci Capitol que te puede interesar |
| a: | is that hotel near palacio real. | es ese hotel cerca de palacio real. |
| b: | Yes, this 10 minute walk | si, esta a 10 minutos caminando |
| a: | what is the cost. | ¿Cuál es el costo. |
| b: | the cost is 100 euros per room per night | el costo es de 100 euros por habitacion por noche |
| a: | HOW MANY PEOPLE CAN STAY IN ONE ROOM | ¿CUÁNTAS PERSONAS PUEDE PERMANECER EN UNA HABITACIÓN |
| b: | the rooms have unique twin | las habitaciones tienen 2 camas singulares |
| b: | is a bed for one person | es una cama para una persona |
| a: | can only one person stay | Sólo una persona puede permanecer |
| b: | two people can be | dos personas puedes estar |
| a: | does it have a microwave | tiene un horno de microondas |
| b: | If | si |
| a: | is that a yes | Eso es un sí |
| b: | If an if | si es un si |
| b: | Yes, yes has microwave | sí, sí tiene microondas |
| a: | that's great, let's book this room | es genial, a reservar esta habitación |
| b: | ready | listo |
| a: | we're also thinking of watching a football match | También estamos pensando en ver a un partido de fútbol |
| b: | ready, there is a match of real madrid against barcelona on 20 March | listo, hay un partido de real madrid contra barcelona el 20 de marzo |
| a: | I want 4 tickets to that match | 4 Entradas para ver a ese partido |
| b: | well, the total is 6000 euros | bueno, el total son 600 euros |
| a: | that's too expensive, do you have anything cheaper | Eso es demasiado caro, tienes algo más barato |
| b: | Yes, I have tickets for 50 euros, but with limited view | sí, tengo boletos por 50 euros pero con vista limitada |
| a: | that sounds ok, i'd like these seats | Eso suena bien, me gustaría que estos asientos |
| a: | what is the location of these seats | ¿Cuál es la situación de estos asientos |
| b: | what you mean with the situation | a que te refieres con la situacion |
| a: | i said location | Dije ubicación |
| b: | the location is good, row H at the top | bueno, la ubicacion es fila H en la parte superior |
| a: | ok, book it | bien, la reserva |
| b: | give me your credit card number | me da su numbero de tarjeta de credito |
| a: | the number is 451234567000 | el número es 451234567000 |
| a: | 45623480809 | 45623480809 |
| a: | one more thing, how do i get to the stadium from my hotel room | una cosa más, cómo al estadio desde mi habitación de hotel |
| b: | You can take a taxi, ask taxi by cuesta de san vicente | puedes tomar un taxi, preguntale al taxi por cuesta de san vicente |
| a: | we are spending a week there, and want to go have some fun. What would you recommend | Nos están pasando una semana allí y queremos divertirnos. ¿Qué les recomendarías |
| b: | many like to go to watch football, especially there is a match of real madrid against barcelona on March 20. Also there are many restaurants in the area | a muchos les gusta ir a ver futbol, especialmente hay un partido de real madrid contra barcelona en marzo 20. Tambien hay muchos restaurantes en el area |
| b: | There is a good Museum of modern art also | hay un buen museo de arte moderno tambien |
| a: | we would like to try some authentic spanish food, is there a restaurant that is not expensive? | ¿Nos gustaría probar la auténtica comida española, hay un restaurante que no es caro? |
| b: | Yes, there is a restaurant called Gaudí that is very popular and cheap | sí, hay restaurante que se llama Gaudi que es muy popular y es economico |
| a: | ok, do I need to book a table there? Or kind of just walk in to those restaurants | bien, ¿tengo que reservar una mesa? O simplemente entrar a esos restaurantes |
| b: | Gaudi does not take reservations, but it has to be with time because it is very popular. I recommend you reach the restaurant 10 hours earlier | gaudi no toma reservas pero tiene que estar con tiempo porque es muy popular. Le recomiendo llegar al restaurante 10 horas antes |
| b: | no forgiveness, 1 hour | no perdón, 1 hora |
| a: | yes I will go there, thanks for the tip | Sí iré allí, gracias por el dato |
| b: | done, with pleasure | listo, con mucho gusto |
| a: | is there also some bar that we can go to have a few drinks afterwards | También hay un bar que podemos ir a tomar unas copas después |
| b: | yes there is a bar called candelaria is very popular | si hay un bar que se llama candelaria que es muy popular |
| a: | can i walk to the bar from the restaurant | Puedo caminar a la barra del restaurante |
| b: | No, you need a taxi unless you want to walk 3 km | no, necesita un taxi a menos que quiera caminar por 3 kilometros |
| a: | yea 3 km is fine, i might just walk to it is the area safe | Sí 3 km está bien, tal vez sólo a pie que es la zona segura |
| b: | If the area is secure but I advise you to be careful and watch | si el area es segura pero le aconsejo que tenga cuidado y este atento |
| a: | sure, thank you | Sí, gracias |
| a: | I'm looking forward to my trip thank you for your help | Estoy deseando mi viaje gracias por tu ayuda |
| b: | Thank you was a pleasure to serve you | gracias fue un placer servirle |

## Test Subject #1: Jenny

|  |  |  |
| --- | --- | --- |
| User | English | Spanish |
| Jenny: | Hi, my name is Jenny and I was looking to book a trip to Spain on March 16 to March 22 | Hola, mi nombre es Jenny y estaba tratando de reservar un viaje a España el 16 de marzo al 22 de marzo |
| Dan: | well, I'm going to help. There is a hotel called Vincci Capitol that may interest you | bueno, yo te voy a ayudar. Hay un hotel que se llama Vincci Capitol que te puede interesar |
| Jenny: | is the Vinci Capitol near the Palacio Real, what's the price range | es el Capitolio de Vinci cerca del Palacio Real, ¿cuál es el rango de precios |
| Dan: | Yes, it is a 10 minute walk | si, queda a 10 minutos caminando |
| Dan: | the price is 100 euros per night | el precio es de 100 euros por noche |
| Jenny: | that sounds pretty good, uh, I'm actually looking for a 2 queens and a view of the street, do you know if there are bedrooms like that open | Me parece muy bien, eh, en realidad busco un 2 reinas y una vista de la calle, sabes si hay dormitorios así abierta |
| Dan: | Yes, all rooms are 2 single beds | si, todas las habitaciones son de 2 camas sencillas |
| Jenny: | is it ok | ¿está bien |
| Jenny: | ok that sounds really good | ¿Suena muy bien |
| Jenny: | while there I would like to see the baseball match, and curious about teams and information | mientras que allí me gustaría ver el béisbol coinciden y curiosidad de los adolescentes y la información |
| Dan: | Baseball is not popular here | baseball no es muy popular aqui |
| Jenny: | oops, I'm so sorry, I'm an american. Football matches happening in the area | Uy, lo siento, soy un americano. Partidos de fútbol en el área |
| Dan: | sure, there is a match of real madrid against barcelona on March 20 | seguro, hay un partido del real madrid contra barcelona en marzo 20 |
| Jenny: | great, ok how much are the tickets and how do i book them | Genial, muy bien cuántas son las entradas y cómo puedo reservarlos |
| Dan: | the tiqutes are worth 10000 euros and I can help you book them | los tiqutes valen 10000 euros y yo te puedo ayudar a reservarlos |
| Jenny: | ok, that sounds good. I would like to book them for that day and how to get to the station. and I need for tickets | Bueno, eso suena bien. Me gustaría reservar para ese día y cómo llegar a la estación. y necesito para entradas |
| Dan: | Sorry, needs few tickets | perdon, necesita cuantos tiquetes |
| Jenny: | i think that is 4 tickets as in the number | Creo que es 4 entradas como en el número |
| Dan: | perfect to go to the stadium you can take a taxi and ask for slope of vicente | perfecto, para ir al estadio puedes tomar un taxi, y preguntar por cuesta de vicente |
| Jenny: | can you define what slope of vicente is, is that where the stadium is | ¿puedes definir lo que cuesta de vicente es que donde está el estadio |
| Dan: | If a site is tourist remaining Stadium certa | si es un sitio turistico que queda certa del estadio |
| Jenny: | is it ok if I get your friend | ¿está bien si entiendo tu amigo |
| Jenny: | I'm so sorry I think I misspoke. I need your friend in case I need to call again | Lo siento, que creo que me equivoqué. Necesito a tu amigo en caso de que tengo que llamar otra vez |
| Dan: | do you need to be my friend? | necesitas ser mi amigo? |
| Jenny: | oh so sorry, I'm having a hard time talking today, I need your name | ¡ lo siento, estoy teniendo un mal rato hablando hoy, necesito tu nombre |
| Dan: | no problem, my name is Daniel | no hay problema, mi nombre es Daniel |
| Jenny: | thank you, I would like to confirm that booking | Gracias, me gustaría confirmar la reserva |
| Dan: | Give me your credit card | Me da su tarjeta de credito |
| Jenny: | Ok, the number 4535124553251804 | Bien, el número 4535124553251804 |
| Dan: | ready thank you very much, your bag was accepted | listo muchas gracias, Su tarjeta fue aceptada |
| Jenny: | I'm sorry I'm having some trouble with my computer system. My bag was accepted? | Siento que estoy teniendo algunos problemas con mi ordenador. ¿El bolso fue aceptado? |
| Dan: | Sorry, your credit card was accepted | perdon, su tarjeta de credito fue aceptada |
| Dan: | It was a pleasure to help you | fue un placer ayudarle |

## Test Subject #2: Katie

|  |  |  |
| --- | --- | --- |
| User | English | Spanish |
| Katie: | Hi, I'm Katie | Hola, soy Katie |
| Katie: | I would like to book a room for 5 people near the Palacio real we will be visting March 16 22 and cost | Me gustaría reservar una habitación para 5 personas cerca del Palacio real que será visitar el 16 de marzo 22 y costo |
| Dan: | Hello, my name is Daniel and I would like to help | Hola, mi nombre es Daniel y me gustaria ayudarte |
| Dan: | There is a hotel called Vincci Capitol that may interest you | Hay un hotel que se llama Vincci Capitol que te puede interesar |
| Katie: | How many brooms will it be, and like single bedrooms | Escobas cuántos va a ser y como dormitorios individuales |
| Dan: | do ready, all rooms have 2 single beds, okay? | listo, todas las habitaciones tienes 2 camas sencillas, esta bien? |
| Katie: | ok yea, how much would the total cost be | Bueno sí, cuánto el costo total sería |
| Dan: | the cost would be of per night | el costo seria de por noche |
| Katie: | ok, how much would the cost be for 6 wait nights 7 | bien, cuánto el costo sería para esperar de 6 noches 7 |
| Katie: | how much would the cost be for the 6 nights | ¿Cuánto el costo sería para las 6 noches |
| Dan: | ready the cost serious tota of 600 euros | listo el costo tota seria de 600 euros |
| Katie: | that would be fine | Eso estaría bien |
| Dan: | more than you want? | algo mas que le interese? |
| Katie: | is there a pool in the hotel? | ¿Hay una piscina en el hotel? |
| Dan: | If there is a nice swimming pool with bar | si hay un piscina muy agradable con bar |
| Katie: | ok, that's all. | Bueno, eso es todo. |
| Dan: | well, something else to help you? | bueno, algo mas que le pueda ayudar? |
| Katie: | could you give us directions to the football stadium? | ¿Nos darías indicaciones para el estadio de fútbol? |
| Dan: | ready, near san vicente slope is football stadium, you can use a taxi to get there and tell him to avoid the traffic, would be how 3000 euro | listo, el estadio de futbol queda cerca de cuesta de san vicente, puedes utilizar un taxi para llegar alla y le dices que evite el trafico, |
| Katie: | do you know which teams are playing? | ¿Sabes que los equipos están jugando? |
| Dan: | Forgiveness does not understand question | Perdon no entendi la pregunta |
| Katie: | Um, do you know which teams are playing in the football match? | ¿Sabes que los equipos están jugando en el partido de fútbol? |
| Dan: | Of course, they will play real madrid vs barcelona. It will be a great match | Claro, van a jugar real madrid contra barcelona. Va a ser un gran partido |
| Katie: | Do you take master card? | ¿Aceptas tarjeta master? |
| Dan: | Of course, which is the number of your card? | Por supuesto, cual es el numero de su tarjeta? |
| Katie: | The number is 4545 | El número es 4545 |
| Katie: | 4535 | 4535 |
| Dan: | Perfect your bag was accepted | Perfecto su tarjeta fue aceptada |
| Katie: | Thank you | Gracias |
| Dan: | Gladly, it was a pleasure to help you. | Con mucho gusto, fue un placer ayudarle. |

## Test Subject #3: Louis

|  |  |  |
| --- | --- | --- |
| User | English | Korean |
| Louis: | Hi my name is Louis may I have your name? | 안녕하세요 제 이름은 에드워드입니다 귀하의 이름을 할 수 있습니다? |
| Kenny: | Yes, hi. I called Kenny. | 네 안녕하세요. 저는 Kenny 라고 합니다. |
| Louis: | No I am called Louis. What are you called? | 아니 난 에드워드 전화 해요. 무엇 이라고? |
| Kenny: | My name is Kenny. | 제 이름은 Kenny 입니다. |
| Kenny: | What can I help? | 무엇을 도와 드릴까요? |
| Louis: | Hello Kenny. I am trying to book a hotel for me and my friends near Seoul. And I need to know what the price and amenties would be for that | 안녕하세요 케니. 나 서울 근처 내 친구에 대 한 호텔을 예약 하려고 합니다. 가격과 amenties 될 것 이라고 알고 필요가 4는 |
| Kenny: | Oh indeed. | 아 그렇군요. |
| Kenny: | Gangnam near the hotel. | 강남 근처에 호텔이 하나 있습니다. |
| Kenny: | You need to know how many rooms? | 당신은 몇개의 방이 필요하죠? |
| Louis: | I need to know how much it would cost for 4 rooms near seoul. | 나 서울 근처 4 실 비용이 얼마나 알고 있어야 합니다. |
| Kenny: | Each room is $ 100 a night. Room and bed are one. | 각 방은 하룻밤에 100 달러 입니다. 방에 침대 하나 씩 있습니다. |
| Louis: | Do you know what comes with the rooms like microwave or refridgerators? | 당신이 무슨 전자 레인지 또는 풍기 같은 객실 함께 제공 알아? |
| Kenny: | Your room is equipped with a refrigerator and a microwave. | 네 방에 냉장고 하고 전자 레인지가 구비되어 있습니다. |
| Louis: | That sounds good I would like to stay at that hotel please. | 제발 그 호텔에서 숙박 하 고 싶습니다 좋은 소리. |
| Kenny: | Yes, I see. Besides it help you? | 네 알겠습니다. 그것 외에 도와드릴게 있나요? |
| Louis: | Yeah, I was hoping I could see a K pop concert as well. Do you know any artists and how they would charge? | 네, 케이 팝 협주곡 ashell를 볼 수 있었으면. 모든 예술가 그들이 어떻게 요금을 것 이라고 아십니까? |
| Kenny: | Oh Kpop enthusiasts. | 아 Kpop을 좋아하시는구나. |
| Kenny: | Yes, that happened to psy concerts. The price is $ 50. | 네 그날 마침 psy 콘서트가 있습니다. 가격은 50불입니다. |
| Louis: | that sounds good. Do you know where the concert is located in? | 좋은 소리. 콘서트에서 어디 세요? |
| Kenny: | Yes, the concert hall is a 10-minute walk from the hotel. | 네 콘서트 홀은 호텔에서 10분 거리입니다. |
| Kenny: | Coex mall and find. | Coex mall 을 찾아 가시면 됩니다. |
| Louis: | So with the hotel rooms and the concert for me and my friends, how much is the trip going to cost? | 그래서 호텔 객실과 나 내 친구에 대 한 콘서트를 얼마나 많은 여행 것입니다 비용? |
| Kenny: | Tell your friends about this for a few minutes? | 친구분이 몇분이시죠? |
| Louis: | No how much is the trip going to cost? | 얼마나 더는 여행 비용 거 야? |
| Kenny: | How many friends? | 친구가 몇 명입니까? |
| Louis: | 4 | 4 |
| Kenny: | Altogether 750 dollars. | 다 합해서 750달러 입니다. |
| Louis: | Alright, that's fine. I would like to pay for this booking. | 좋아, 그건 괜 찮 아 요입니다. 이 예약에 대 한 지불 하 고 싶습니다. |
| Kenny: | Yes, I see. Please tell me a credit card number. | 네 알겠습니다. 신용카드 번호를 말해 주십시오. |
| Louis: | My credit card number is four five three five. | 내 신용 카드 번호는 4 5 세 5. |
| Kenny: | 4535? | 4535 입니까? |
| Louis: | yes | 예 |
| Kenny: | Yes, thank you. Fully booked. | 네 감사합니다. 예약이 완료 되었습니다. |
| Louis: | Thank you. | 감사합니다. |
| Kenny: | Yes, thank you. | 네 감사합니다. |