

The Panlingual Camera Phone: Device-Based Pilot Usability Study

Martin Hecko, Kinsley Ogunmola, Jonathan Pool, Tim Wong, Peter Woodman

Project sites:

http://www.cs.washington.edu/education/courses/cse490f/07wi/project_files/camera/
<http://panlingual.org/> (hosted by [Utilika Foundation](#))

Report URLs:

http://www.cs.washington.edu/education/courses/cse490f/07wi/project_files/camera/rereproto/
<http://panlingual.org/rereproto/>

Introduction

We report here a pilot usability study of a [device-based prototype](#) of the user interface for a [panlingual camera phone](#).

The panlingual camera phone will be an application running on a camera-equipped mobile telephone. The application will support the task of understanding written and printed texts in all languages everywhere in the world. A person who sees a text, such as a street sign, protest placard, movie title, or restaurant menu, can photograph the text with the camera and request a translation of it. The application utilizes user preferences, on-device functionalities, remote servers, and human service providers to (1) identify the original language, (2) recognize the text in the image, (3) determine the part of the text to be translated, (4) determine the language into which to translate the text, (5) translate the text, and (6) display the translation to the user. Steps 2 and 3 can optionally be bypassed when the user enters text manually or selects text from an existing document. An optional alternate input source is any existing image that contains text.

After developing and testing low- and medium-fidelity prototypes on paper and computer monitors, we created a prototype of the user interface on an illustrative mobile telephone. The new prototype incorporated some revisions based on prior user testing. It also was more realistic than its predecessors because it ran on an actual mobile device. This prototype, unlike its predecessors, permitted users to take real photographs, select photographs from a library for translation, enter text as input for translation, and check and correct the application's recognition of text found in images. After this increment in prototyping, the interface was ready for additional user testing. The experiment we report here was a pilot study for such an experiment. Its purpose was to test and revise our procedure before we conduct a study with more participants.

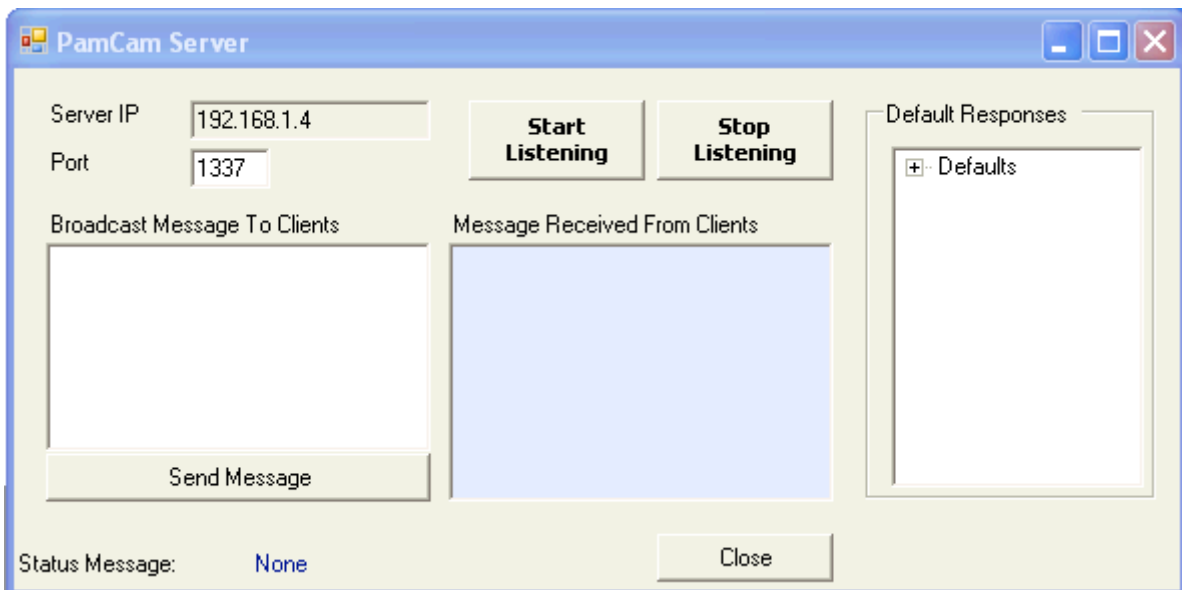
Method

Participants

We conducted the study on five participants, selected as a convenience sample from acquaintances and strangers encountered in a student computer laboratory. All five participants were students. Their ages ranged from 19 to 23 years. All reported using mobile telephones, and all reported they had experience taking photographs with mobile telephones. Obviously, our sample was less heterogeneous than it should have been, but nonetheless, as we shall describe, exhibited substantial variation in performance success. The detailed participant data are in Appendix 3.

Apparatus

The client device was a Cingular 8125 mobile communication device, equipped with a 1.3-megapixel camera, a 39-key alphanumeric keyboard, a 320 x 240 color touch-sensitive display, a stylus, and multi-protocol communication. Its operating system was Windows Mobile 5.0. We simulated the on-device services, remote server, and remote human service providers with a server application running on a portable computer operated by one of the experimenters.



Tasks

We instructed our participants to attempt to perform three tasks with the device. The instructions were as follows:

Task 1: Understanding something around you. Suppose you want to understand what a sign says. This device will let you take a picture of that sign and get a translation into English. Please try this now.

Task 2: Making yourself understood. Suppose you want to give a note to somebody near you who can understand only Hattanese. Think of something to say ("Is there a good restaurant around here?" or whatever else you want). With this device, you can enter your message in English and get it translated into Hattanese, so you can show it to the other person. Please do that now.

Task 3: Understanding a scene you saw earlier. Suppose you want to choose a photograph from your library and understand the text in it. This device will let you choose a photo, and it will show you

how it reads the text in the original language. If you see any error, you can correct it. Then you can have the corrected text translated into English. Now please choose a photo, correct any errors you see in the reading, and then get it translated into English.

Procedure

1. We explained the purpose of the study to the prospective participant, answered any questions, and, if the prospect agreed to participate, obtained the participant's consent with a written and signed participation agreement.
2. We invited the participant to complete a brief questionnaire asking about some personal facts and some indicators of the participant's experience with mobile information technology.
3. We read to the participant an explanation of the participant's role in the study, emphasizing that the participant was not being evaluated, but was helping us evaluate the design of our user interface, and inviting the participant to think aloud so we could better understand what parts of the interface worked well and worked poorly.
4. We demonstrated some of the use of the Cingular 8125 device. In particular, we showed the participant that the device has a slide-out keyboard and a stylus. Thus, we did **not** show the participant how to use the device's hard buttons labeled at the bottom of the display, how to take a photograph, or how to enter text with either the hardware keyboard or the on-display keyboard. We told the participant that some features of the interface weren't working yet and would be faked in this study. We invited the participant to ask any questions based on this demonstration, and we answered them.
5. We read the instructions for the tasks to the participant. After each instruction, we waited, observed, and video recorded while the participant attempted to perform the task. One of the experimenters operated the server and performed text transmissions to the participant's device as necessary. If the participant got stuck and could not proceed without help, one of the experimenters provided enough help to let the participant continue performing the task. The experimenters made notes of critical incidents, including features that performed notably well and serious problems.
6. After the final task, we invited the participant to complete a brief final questionnaire and undergo a short oral interview. The questionnaire asked the participant to rate the user interface's ease of use and understandability. In the interview, we asked the participant for comments and suggestions. Finally, we invited the participant to ask any questions and answered them.

Test Measures

We measured relevant participant experience with the participant's answers to our questions on the participant's familiarity with mobile information devices and knowledge of multiple languages. We wanted to know whether more experienced participants would more easily use our interface and whether some minimum experience would be a prerequisite to success.

We measured task performance success by recording for each participant and each task whether the participant completed the task, how many times the participant needed help, how many errors the participant made, and the duration of the task. We wanted to know whether our interface was usable and which functionalities were difficult.

We measured participant satisfaction by obtaining the participant's answers to our questions on how easy and understandable the interface was. We wanted to know how users would feel about our interface and how their subjective response to it would be related to their objective performance using it.

We also recorded observations of participants' oral comments and participant successes and failures in the use of the prototype. We wanted to capture any important design-relevant facts that we might notice, for use in further development.

Results

Task Duration

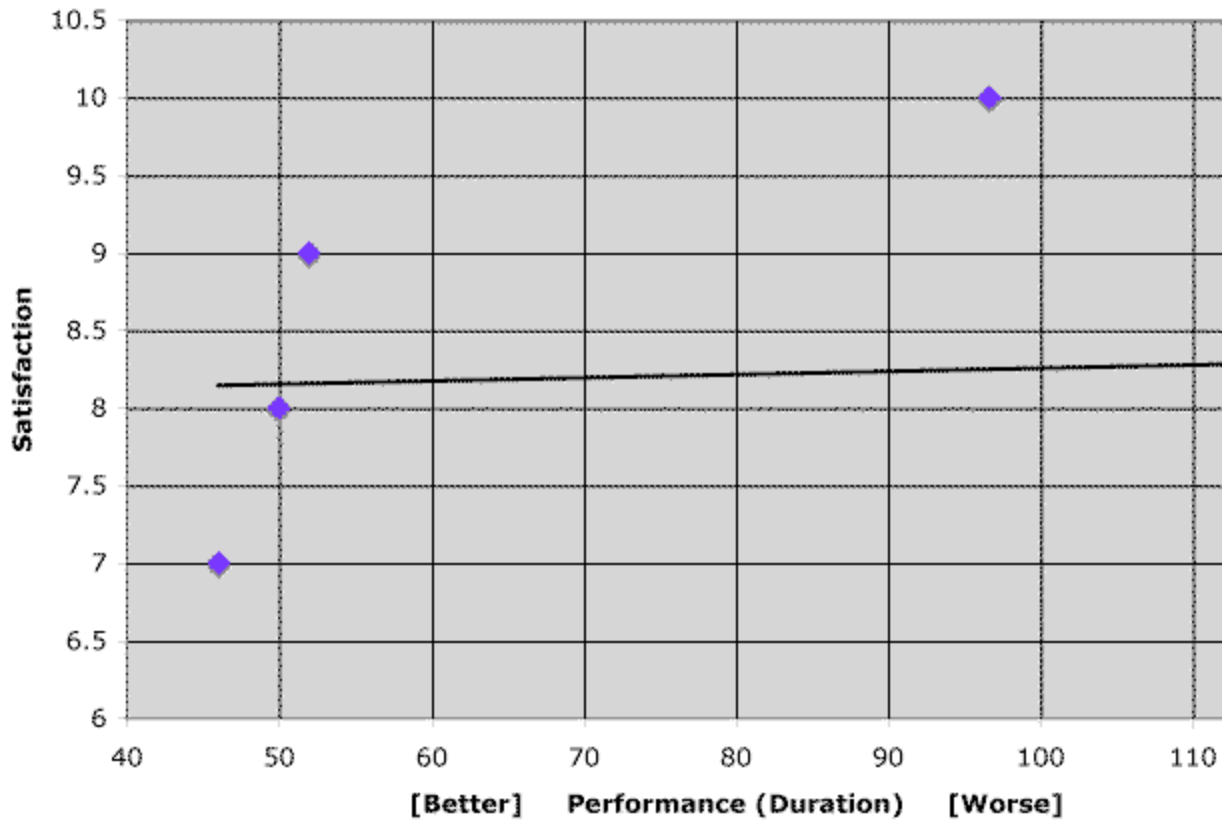
The task durations in seconds were as follows:

Participant	Task 1	Task 2	Task 3	Mean
1	53	45	52	50
2	44	47	65	52
3	49	59	30	46
4	20	90	180	97
5	115	176	63	118
Mean	56	83	78	76

Performance and Satisfaction

We expected participant satisfaction to covary positively with participant success in task performance. The results failed to show any such covariation. Satisfaction is measured here as the sum of the ease-of-use and understandability scores, where "excellent" and "good" understandability are recoded as 5 and 4, respectively.

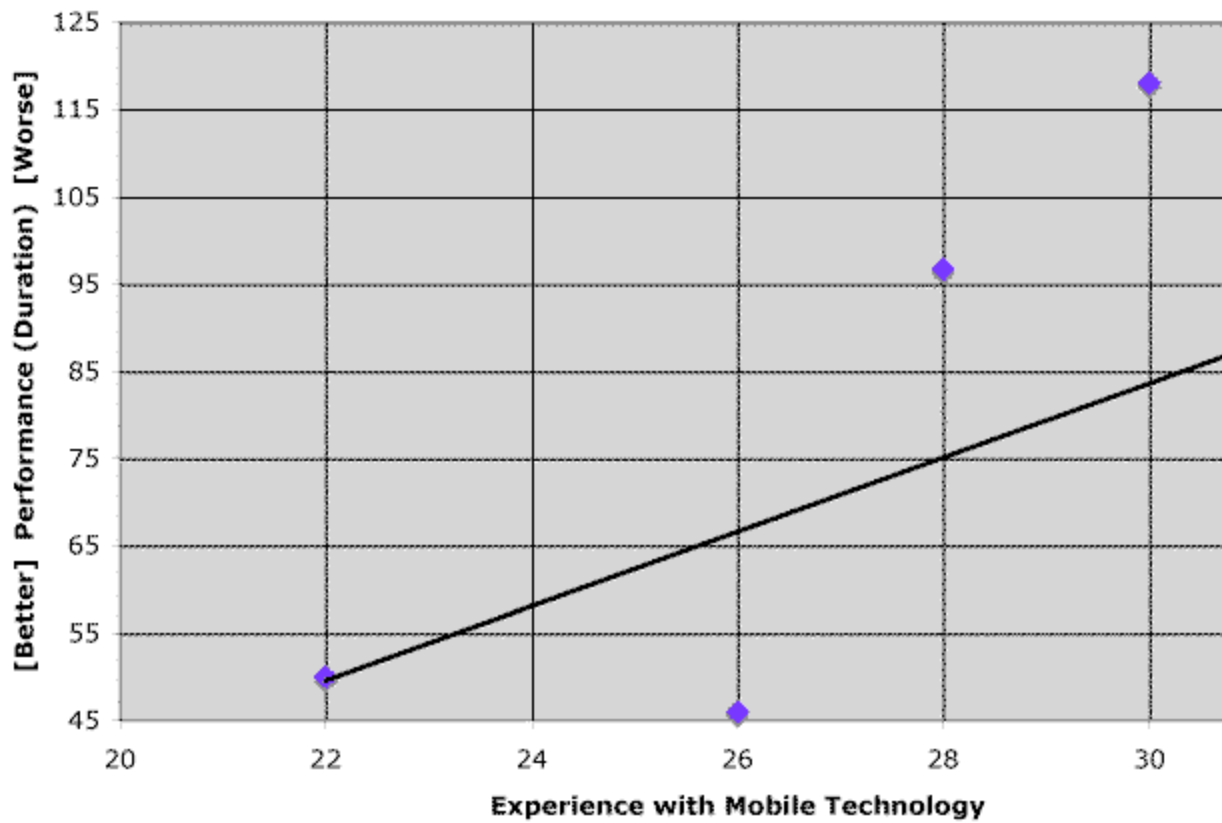
Performance and Satisfaction



Mobile Technology Experience and Performance

We expected participant performance to covary positively with participant mobile technology experience. The results failed to show any such covariation. Instead, they showed a negative covariation. Experience is measured here as the sum of all device-use and mobile-phone-task-ease scores, where device-use responses of "yes" and "no" are recoded as 4 and 2, respectively.

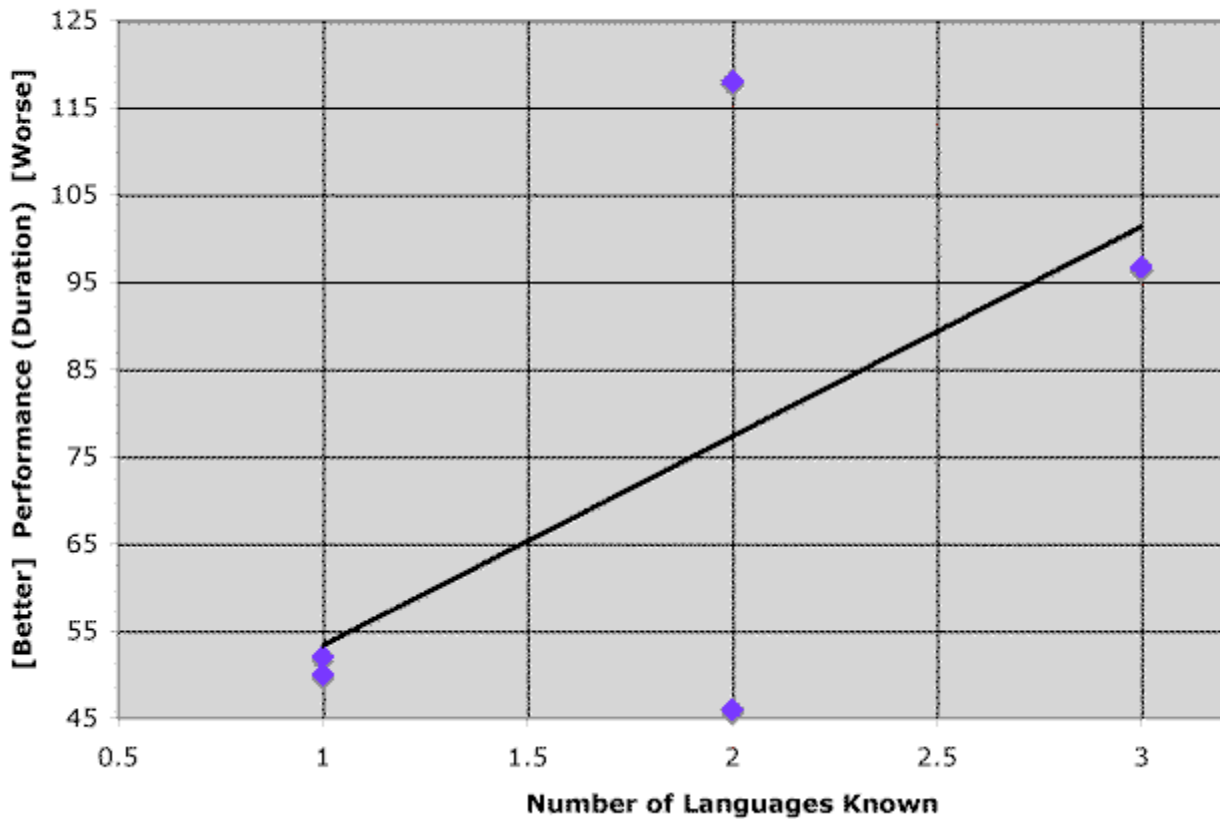
Mobile Technology Experience and Performance



Language Knowledge and Performance

We expected participant performance to covary positively with participant multilingual skill. People knowing multiple languages would tend to understand the purpose of the device and better distinguish OCR output from translation output. The results failed to show any such covariation. Instead, they showed a negative covariation.

Language Knowledge and Performance



Errors and Comments

The fifteen task trials (five participants, with three tasks per participant) produced six critical incidents. Of these, two had severity 4 (critical), four had severity 3 (serious), and none were observations of notably good features. Most incidents appeared to be avoidable with better help. However, the critical incidents were cases in which the prototype could not handle particular button presses on the device, with the effect of task-performance failure and/or loss of previously entered data.

Participant comments included praise for the interface, complaints, and recommendations. Participant satisfaction was in all cases medium to high.

Discussion

The pilot study delivered results justifying the pilot-study method. The study was fast and inexpensive, but revealed some serious interface deficiencies that merit correction before a larger-scale study is conducted. This is despite the generally high level of experience with mobile technology exhibited by the pilot-study participants. However, in a larger-scale study it appears essential to recruit participants with a substantially wider variety of technology experiences and life situations than we recruited here.

The pilot study also seems to have shown us that demonstrations of the features of the underlying device should be more thorough in a larger-scale study. Most problems encountered by our participants were at least partly due to unintuitive features of the device, rather than our application.

Observed errors and participant comments led to some ideas about improvements in the interface design. The main such ideas are:

- Making more help available (this is limited when the underlying device's affordances are relied on but are deficient in usability).
- Assuring that erroneous button presses can be recovered from with no lost user input.
- Providing options to move from mode to mode, e.g. from photo input to text input.

Appendix 1: Materials

Invitation

We are students in a computer science course at the UW, and we're developing an application that would run on a mobile phone. It would let people take pictures of signs, menus, and other written and printed things with their phone cameras. Then it would get the texts in those pictures translated, so people could understand what is around them anywhere in the world. As an additional service, it could get translations for other texts, including texts typed in by the user.

As we develop this application, we are evaluating it by asking some anonymous volunteers to try using it. Would you be willing to help us by trying it out now? It would take about 15 minutes.

Participation Agreement

Participation Agreement

Information for Participant

We are students in a Computer Science course, CSE 490F, at the University of Washington. As part of our work in this course, we are developing a Panlingual Camera Phone application. Periodically, we evaluate it by asking some volunteers to try using it. The volunteers provide data that we use as we further develop the application. We collect data by interview, observation, questionnaire, and video tape.

Your participation in this study is voluntary. You may withdraw yourself and your data at any time without fear of consequences. You may discuss any concerns about the study with us (Martin Hecko, Kinsley Ogunmola, Jonathan Pool, Tim Wong, and Peter Woodman), or with Professor James A. Landay, the instructor of CSE 490F. Professor Landay may be contacted by telephone at 206-685-9139 or by Email at landay@cs.washington.edu.

Your participation in this study will be anonymous. Your name will not be asked or recorded. You will sign both copies of this sheet with an anonymous identifier of your choice.

Agreement to Participate

I hereby acknowledge that I have been given an opportunity to ask questions about the nature of the study and my participation in it. I agree to have data collected on my behavior and opinions in relation to the Panlingual Camera Phone study. I understand I may withdraw my agreement at any time.

Anonymous identifier _____

Date

22 February 2007

Witness name _____

Witness signature _____

Orientation

We're going to show you a couple of things about this device first. Then we'll ask you to try to perform three tasks with it. While you're trying, we'll want you to think aloud, so we can understand what's easy, what's hard, and what's impossible, and why. Of course, we're testing our user interface here. It certainly has defects, and we want to find them. We're not testing you. If you can't perform the tasks, it means our interface needs to be improved, and you'll be helping us to understand what's wrong with it. Do you have any questions before we show you how the device works?

Initial Questionnaire

Panlingual Camera Phone: Anonymous Participant Questionnaire

Your sex: Male / Female

Your age: _____

Your occupation: _____

Which of the following do you use? (Check all that apply.)

- Cell phone
- PDA (Palm)
- Portable music player (Zune, iPod, etc.)
- Smart phone (Blackberry, Palm Treo, Windows Mobile phone)

On a scale of 1 to 5, 1 being the least comfortable and 5 being very comfortable, rate how comfortable you are using the following features of your phone:

	1	2	3	4	5
taking a picture	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
browsing the Internet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sending text messages	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sending email	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sending a picture message	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How many languages, including English, do you know well enough to read a newspaper with help from a dictionary? ____

Test

Now we're going to read you instructions for three tasks. After we read each instruction, you'll try to do what we said, using the device. Ready? Any questions before we start?"

[Task instruction 1]

Great. Thank you. Now the next task.

[Task instruction 2]

Great. Thank you. Now the final task.

[Task instruction 3]

Great. Thank you. That's the end of the tasks. Now we want to finish by getting a little information about you and your impressions of this device.

Questionnaire-Based Survey

**Panlingual Camera Phone:
Final Questionnaire**

Overall please rate the user interface and usability of the device:

How easy did you find this device to use to accomplish the tasks? (Please circle: 1 = worst / 5 = best.)

1 2 3 4 5

How understandable was the user interface to use to complete tasks?

Excellent / Good / Mediocre / Poor / Horrible

Thank you very much for your help!

Appendix 2: Critical Incident Logs

Task	Category	Description
1	Bad 3	P5 had much trouble with the Windows Mobile default camera interface. He got confused by the many options and the confirm screen.

2	Bad 3	P1 attempted to use the photo translation screen to translate text, which it doesn't support. This should be rectified in the future. It caused the participant to get stuck. To resolve the error we advised the participant to go to the main screen and start again.
2	Bad 3	P2 was unsure how to enter text into the device. The stylus had to be pointed out to P2.
2	Bad 4	P4 kept exiting to the menu in text translation. P4 was operating in keyboard-slideout mode, which our interface does not adapt to. P4 was pressing "enter" and triggering the "back" default action. This should be corrected.
3	Bad 4	P1 accidentally hit the back button, which led to the main menu. P1 had been editing a text field. When P1 got back to the editing screen, the edited text had vanished and been replaced by the original text.
3	Bad 3	P1 didn't know how to return to the picture selection screen from the picture translation screen.

Note. Categories are "Bad 4" (critical), "Bad 3" (Serious), and "Good".

Appendix 3: Questionnaire Responses

Participant	1	2	3	4	5
Sex	Male	Male	Female	Female	Male
Age	20	21	23	22	19
Occupation	Student	Student	Student	Student	Student
Uses Cell Phone	Yes	Yes	Yes	Yes	Yes
Uses PDA	No	No	No	Yes	No
Uses Music Player	Yes	Yes	Yes	Yes	No
Uses Smartphone	No	No	No	No	No
Taking Picture Ease	5	3	4	4	4
Browsing Net Ease		3	1	2	4
Text Messaging Ease	2	5	5	3	4
Email Sending Ease		3	1	2	4
Pic Messaging Ease		5	3	3	4
Languages Known	1	1	2	3	2
UI Ease of Use	4	4	3	5	3
UI Understandability	Good	Excellent	Good	Excellent	Good

Note. Numeric codes for ease range from 1 (least easy) to 5 (easiest).

Appendix 4: Participant Comments

Participant 1

Device Intuitive?

What to do in the beginning. Used to the top being menu

Overall comments:

No easy way to go back. Keyboard covers up the ____ text, when the physical keyboard is pulled out, no scrollbar

From task 1 to task 2, problem with _____

Participant 3

Device Intuitive?

Doesn't know when to click

Intro paragraph would be helpful

Doesn't know how to use the windows mobile device

Participant 4

Note: didn't try to correct the text

Device Intuitive?

Took a while to translate

Don't really know what to click when looking at picture

It would be better if there is like a description or intro paragraph that says this is step by step if you wanna do something this is the step that you should follow

That would be really helpful

I'm not really familiar with this device, so I had a hard time figuring out what's supposed to be what

Program is pretty neat

Participant 5

Device hard to find out that middle button takes picture

Keyboard easy to use.

Get stuck in a "hole" when taking a picture in the photo menu

