Context-Sensitive Microlearning of Foreign Language Vocabulary on a Mobile Device

Jennifer S. Beaudin[1], Stephen S. Intille[1], Emmanuel Munguia Tapia[1], Randy Rockinson[1], Margaret E. Morris[2]

[1] House_n, Massachusetts Institute of Technology
One Cambridge Center, 4FL, Cambridge, MA 02142 USA
jbeaudin | intille @mit.edu
[2] Digital Health Group, Intel Corporation,
20270 NW AmberGlen Court; AG1-102, Beaverton OR 97006 USA
Margaret.Morris@intel.com

Abstract. We explore the use of ubiquitous sensing in the home for context-sensitive microlearning. To assess how users would respond to frequent and brief learning interactions tied to context, a sensor-triggered mobile phone application was developed, with foreign language vocabulary as the learning domain. A married couple used the system in a home environment, during the course of everyday activities, for a four-week study period. Built-in and stick-on multi-modal sensors detected the participants' interactions with hundreds of objects, furniture, and appliances. Sensor activations triggered the audio presentation of English and Spanish phrases associated with object use. Phrases were presented on average 57 times an hour; this intense interaction was found to be acceptable even after extended use. Based on interview feedback, we consider design attributes that may have reduced the interruption burden and helped sustain user interest, and which may be applicable to other context-sensitive, always-on systems.

Keywords: microlearning, language learning, context-sensitive, context-triggered, mobile phone, sensors, home, deployments

1 Introduction

In this work we explore the use of context-sensitive ubiquitous computing for learning. Although many applications have been proposed that use ubiquitous computing and context-aware sensing for reminders, entertainment, medical monitoring, and communication systems, relatively little work has been reported on how context-aware systems might be exploited to build applications that help people incrementally learn new things. Yet, automatic detection of context may enable learning tools that present information and interactions to people at appropriate moments as they engage in their everyday lives. We describe our experience building and testing such a system, and we report some observations that may be relevant to those interested in creating other types of "always-on" applications that use ubiquitous computing and automatic context detection.

The inspiration for this work was an eLearning technique called “microlearning,” where a difficult learning task is broken into a series of very quick learning interactions, distributed over time [1]. Rather than having to learn or practice everything at once, the learner is presented with a small, manageable chunk of information at regular intervals. Researchers in previous work have explored how to deliver learning interactions using technology during moments when the user may be more receptive, in idle moments such as during the start-up period of computing devices and when computer screen-savers are automatically initiated [2].

The brief interactions used by microlearning systems may not be appropriate for in depth learning, but they may allow users to chip away at a larger learning goal. Moreover, temporally-spaced presentation has been shown to yield higher learning rates than massed presentation (i.e., “cramming”) [3]. Microlearning interactions may also serve a “priming” role by repeatedly bringing the learning task to the user’s attention; users may then be more mentally prepared to take advantage of richer learning opportunities, such as those that occur in classrooms and in natural domain-interaction contexts.

Ubiquitous technology can be usefully applied for microlearning because it can reach users throughout the day, when they have idle time, and in contexts that are related to the information being learned. Research on “encoding specificity” has suggested that information is memorized in combination with contextual cues that facilitate retrieval in matching contexts ([4, 5]). A person is more likely to recall something if it was learned in association with cues that will be available at the time of retrieval. For example, if a person learns information when she is in a specific physical context (e.g., underwater [6]), she is better able to recall that information later when in that same context. While factors such as task attention requirements affect the extent to which context acts as a retrieval cue [7], these findings suggest that learning information in contexts similar to those where the information will be needed (exploiting ubiquitous computing) may aid later retrieval.

Additionally, learning that occurs without formal instruction, such as first language acquisition, seems to be based on *meaningful* exposure and interaction [8]. That is, learners are most receptive to information, ideas, and skills that are relevant to their current needs and actions. Context detection may provide the opportunity to deliver learning interactions at times when users can immediately apply what they learn.

1.1 Acceptance and Usability of an Always-On Application

We developed this language learning system to explore how ubiquitous technology may enable users to chip away at a larger learning task in a non time-intensive way and during periods when they are better able and more motivated to do so.

The real benefit to users of systems such as these would likely come from frequent engagement over an extended time period. However, data about how such systems, once deployed, are used over time are as yet underreported. Questions that are difficult to answer in the absence of usage data and user feedback include whether acceptance would wane after a period of novelty, what types of impromptu interactions would result, and to what extent users would perceive such a system as valuable, given the sensing requirements and interruption burden. These questions apply to a variety of ubiquitous systems that present information in context.

To begin to address these questions, we conducted a study in which a microlearning prototype was deployed in a real-life context for an extended time. Participants lived with the application as part of their day-to-day home life, and after the study period, they were interviewed. We use this case study to show proof of concept that a ubiquitous system that presents information to users in context at rates as high as 142 times per hour might be tolerable for several weeks of intense use. We also use the case study to gather information about how to improve a system that presents continuous microlearning feedback.

1.2 Learning Domain: Second Language Vocabulary

Second language learning was chosen as our application domain because it affords a straightforward mapping between content and context (i.e., foreign language labels that relate to objects and events), it is a learning task of interest to many people at different stages of life, and it is not easily sustained by learners through other means for an extended period of time.

Bilingualism is essential for many people in the world today and even in countries where citizens do not need to learn more than their native language, it is recognized that learning a second language is valuable for economic and social mobility [9]. For those interested in language learning, however, the challenges of becoming proficient, particularly when not immersed in a culture speaking the language, can be daunting. The Foreign Service Institute estimates that it takes approximately 480 classroom hours to become minimally proficient in a second language that has a similar structure and phoneme system (e.g. English speakers learning Spanish) [10]. It is estimated that a vocabulary of 6,000 to 7,000 words is required for unassisted comprehension of spoken language (analysis for English [11]). While a smaller vocabulary may be sufficient for a traveler's needs, even learning 100 spoken words and being able to recognize them in context represents a significant memorization task.

Adults approach this task using a variety of formal and informal methods, including classroom instruction, books, audio CDs, and multimedia software. These methods can be difficult to sustain, particularly given their time requirements and the motivation issues related to being a novice speaker (i.e., social inhibition). There is a need for methods and technologies that reinforce these more intensive efforts.

A few notable applications of ubiquitous technology for language learning have been proposed and prototyped, including a common-sense language translation application for mobile phones [12] and presentation of foreign language vocabulary incrementally via a screen-saver application [2] and SMS messaging to mobile phones [13]. To our knowledge, using detected context to trigger language-learning interactions has not yet been studied, and we are not aware of any extended deployments of such systems outside of lab settings.

1.3 Context Detection

Context detection for ubiquitous computing is an active area of research, with proposed strategies ranging from using RFID tags on objects to using computer vision for image analysis. Recent work in ubiquitous computing suggests that object-based dense sensing can be used to recognize context, even in complex, real-world settings such as homes (e.g., [14-16]). When using this approach for context detection, miniature sensors are attached directly onto many objects in the environment and can enable a computer system to infer contextual information about the home occupant's movement and everyday activities such as "cooking", "making tea," "vacuuming", and others. These sensors are small and inconspicuous. Further, when mass-produced, they should be possible to manufacture at a low cost. Although other types of context detection could also be used to build context-sensitive microlearning applications, object-based dense sensing was the approach used in the case study we present here. Context was detected by using motion sensors and RFID tags to detect object usage and object touch. Although not used in this experiment, more advanced inference of activity type could also be used in an extension of the proposed system.

One criticism of object-based dense sensing is that the infrastructure, which could consist of tens to hundreds of small sensors stuck on objects, may be too difficult to install and maintain over time. We later address how ubiquitous computing applications that provide immediate value to the end user, such as just-in-time microlearning, might be useful in addressing this concern.

2 System Design

The system was designed to mimic informal strategies that beginning learners of a second language often employ. These include placing vocabulary labels on objects in their home (e.g., taping the word "puerta" to the door), asking friends who are native speakers for words that reference objects or situations in the immediate context (e.g., "How would I ask for a cup of coffee?" while getting coffee), and practicing L1-L2 translation pairs (e.g., "milk - leche").

We wanted to take advantage of the consistency and ubiquity afforded by the sensing and mobile technology - learning can take place anytime and everywhere - while giving the user control over the level of attention and engagement he or she would give to any interaction, including having the option of ignoring it. To do this, we designed the system to be "always on," but to be hands-free with short interactions. To enable users to make progress on a microlearning task, we expected that the application would be run for several weeks or more, yet it needed to sustain interest. The system was designed, therefore, to provide some variability on repeat interactions and to be scalable to exploit improvements in the sensing and inference algorithms.

2.1 Scenario

Here we describe the operation of the system from the user's perspective. Stacy has just arrived home and is wearing her mobile phone at her belt. When she opens the door, she hears "door" (in English) and then the Spanish translation "puerta." She puts her coat in the closet and hears "closet - armario." She sits down on the couch and hears "couch - sofá," followed by "cushion - cojín." "Cojín" is repeated again. She watches television for about 30 minutes, without hearing any phrases. She then gets up to make a snack in the kitchen. She opens the refrigerator and hears "refrigerator - refrigerador." She grabs a container of milk, and hears "milk - leche." That is a word she hasn't heard before, so she pulls out the phone and presses a button to have the phrase repeated. As the phrase is played, she looks at the screen on the phone to see the word. She tries pronouncing the word and then sets the phone on the counter while she continues preparing her meal.

2.2 System Design

The prototype system employed ubiquitous sensing to detect context and a mobile phone for audio and visual presentation. Built-in and wireless stick-on multi-modal sensors detected the participants' interactions with objects, furniture, and appliances and sent signals to a server in real-time. The mobile phone application polled the server using a GPRS connection and responded to sensor activations by showing and audio playing English and Spanish phrases associated with the moved object.

Although the system can be deployed solely with portable sensors, for the purposes of the work reported here, it was deployed in a highly-instrumented home. The live-in home was designed to support the collection of rich, multi-modal sensor datasets of domestic activities and to provide a naturalistic environment for the evaluation of novel ubiquitous computing technologies that use house-wide sensor systems [17]. Volunteers are recruited to move into the instrumented home and treat it as their own home as much as possible during their stay. They often provide explicit feedback during or after the study, which supplements the rich record of their activities and interactions with technologies that is collected using the home's infrastructure. The couple who used the system described here lived in the home for ten weeks and used the language learning tool for the last four.

In this work, four types of sensors were used. The 1000 sq. ft. apartment has eighty small, wired switches embedded discretely in all cabinetry and appliances that were used in this work. The home also has water flow sensors on all hot and cold taps in the unit. In addition, a portable kit of "object usage" wireless sensors was deployed on furniture and other objects in the home (Figure 1a). These sensors contain a sensitive piezo trigger that wakes up an accelerometer when the sensor the object is attached to (or placed in) is moved. The sensors are small and can be attached to most objects with adhesive putty or simply placed in drawers or pockets or under cushions. A sensor ID is transmitted wirelessly to a receiver in the home when the sensor moves.

Figure 1. a) A stuck-on wireless object usage sensor detects movement of the TV remote control; b) the RFID bracelet worn by the user senses “near touch” events, such as when this salt container is retrieved from the pantry; c) A simple UI: the Spanish phrase “lavaplatos” is displayed and played on the mobile phone when the dishwasher door is opened.

The fourth class of sensors used were radio frequency ID (RFID) stickers, cards, and buttons that were placed throughout the home, on cabinet and appliance surfaces, as well as movable objects, such as cooking utensils, cleaning equipment, food containers (Figure 1b), portable electronics, and books or magazines. In this study, the male volunteer wore an RFID reader built into a bracelet form factor [18]. Whenever the user's wrist comes within a few inches of an RFID tag, the bracelet wirelessly sends the tag ID to a receiver connected to the home’s infrastructure, and the "near touch" event is recorded. The bracelet was taken off each night and when the volunteer left the apartment and plugged in to charge. Although the bracelet is a bit bulky at first, the volunteer was able to adjust to it and wear it continuously while in the unit for 10 weeks.

Computers in the home, enclosed out of view from the inhabitants, receive data from the sensors. The data are time-stamped and stored locally to disk. Identification numbers from selected sensors are then transmitted to a secure server that can be polled by client applications to trigger sensor-activated interactions.

Identification numbers for the sensors were associated with phrases that describe the object the sensor was stuck on or placed in (i.e., “refrigerator”) and objects typically contained within, if any (i.e., “eggs,” “milk”). In a few cases, simple words that may be related to the sensor-tagged object (i.e., “exit, entrance” for door, “science” for a science textbook) were also associated with a sensor ID. These sensor-phrase associations were stored in XML format on the server and were assembled by both the participants and the researchers.

The participants, as part of a concurrent protocol investigating self-installation of home sensors for context-detection, were asked at the beginning of the study period to place the wireless object usage sensors on furniture, objects, appliances, cabinetry, and other movable parts in the home and to label their placement of the sensors using an application on the mobile phone. The participants entered the room location and selected a label for the host object from a list of 1949 items for each placed sensor. Researchers placed additional object usage sensors and RFID sensors on all objects that were untagged and that could robustly accommodate a sensor without impacting object usability and provided object labels for the wired sensors, RFID sensors, and researcher-placed object motion sensors.

Four hundred phrases were selected from this combined inventory to be used by the language application. The selected phrases were nouns and were translated by a research team member into Spanish (Mexican dialect). Spanish was chosen based on the travel interests of the participants, the availability of a Spanish translator, and the prevalence of Spanish as a language in North America. Examples of English-Spanish pairs that were used include "keyboard - teclado" and "aftershave - loción para afeitar." Eighty-five wired switch sensors, 170 wireless object motion sensors, 431 RFID tags, and 14 wired flow sensors had labels with matching vocabulary.

The language learning software was written in C# for phones with the Windows Mobile operating system. For this study, an AudioVox SMT5600 mobile phone was used. The software application polls the server for sensor activations every second. These activations are mapped to any associated words, which are returned to the phone. If no sensor activity is detected for more than 5 minutes, the software polls the server less frequently (once a minute), to save battery. Due primarily to GPRS connection time, there is a latency of 5-15 seconds between action and word presentation on the mobile phone, although a near instantaneous presentation speed can be achieved on a PC. GPRS was selected over WLAN or Bluetooth to maximize battery life (~6 hours versus ~2 hours of active polling on the mobile phone).

When the phone receives phrases corresponding to a recent action within the home, the software filters the list to reduce repetition: a phrase cannot be played more than 2 times in a minute or 6 times in an hour. This decision was made to reduce the interruption burden and to allow a greater variety of phrases to be played during any given period. A phrase is randomly selected from those that remain. The English phrase is displayed on the mobile phone screen while a sound file of the phrase spoken in English is played. Then the Spanish translation is displayed and played (Figure 1c). The application then visually presents an option to repeat the phrase. The only other interface element is a 30-minute mute option.

The application starts automatically when the phone is turned on and does not require user interaction for normal use. For this study, the mobile phone was used only for this application and not for communication purposes. It was expected that the phone would be carried at the belt or in a pocket or purse some of the time. Given that previous research has shown that people do not necessarily carry their phones with them while at home [19], sound play was made loud enough to be heard within a room and adjacent spaces if the phone was set down.

3 Exploratory Study

The exploratory case study was designed to assess the acceptability and perceived value of the system. To this end, we sought to provide a study context that would be similar to a real-world scenario. Participants (who were not aware they would be testing a language application when they signed up for the study) were provided with vocabulary for a language that they were motivated to learn. In order to avoid introducing undue stress or artificial motivators, we did not set an expectation that they would be tested or have them complete comparison learning tasks. As a result, however, we were only able to conduct a limited evaluation of learning performance.

The participants were recruited from a pool of individuals who had responded to postering, electronic mailing lists, and press articles announcing that subjects were needed to study how to make technology easier to use in the home. The participants were a married couple: a woman, age 31, working in the publishing industry, and a man, age 29, a high school science teacher. Although they both worked in science-related fields, they did not have advanced knowledge of computer science or sensor technology. They were recruited for a primary study to investigate personal health monitoring and were asked to try out several mobile-phone applications during their 10-week stay, but they were not given details about these applications beforehand. Among the reasons they provided for their interest in participating was to *"see how we can use technology to simplify our life, not complicate it."*

Before the study, they were asked to describe their foreign language backgrounds, travel history, and countries where they would like to travel. They both had taken a few years of Spanish in high school (10+ years earlier), but admitted that they were very rusty. They listed 13 countries that they would like to visit, and Spanish is the primary language in three.

For the first six weeks of the study, the couple participated in other protocols. At the beginning of the study, they were asked to place wireless sensors throughout the home using a mobile phone application to register and label their locations. They field tested two other mobile phone applications: a names and faces microlearning application [20] and a brief cognitive assessment game. These applications prompted brief interactions on a timed schedule. Over the entire study, data about their daily activities were recorded by the home's sensing infrastructure. The participants were encouraged to maintain as normal a routine as possible. They went to work, had visitors over, cooked meals, attended to sleep and personal needs, and worked on projects and leisure activities according to their own preferences.

The male participant was asked to carry the phone with him whenever he was at home. He was instructed to turn off the application and plug the phone into a charger near the door whenever he left. It was noted from the collected data that on weekdays, the participants were only running the application in the evening. For the final week of the study, the participants were encouraged to try it in the morning as well. Even though the male participant was the primary user, the female participant could also hear words when she was in the unit. Further, sensors activated by both individuals were triggering the word selection.

Our primary goal was to obtain qualitative feedback on the experience of using the context-sensitive microlearning application in the home for four weeks. The male participant provided feedback about how the application was working four times during the study by email. Both participants were interviewed together after the study in a 90-minute session; the interview was audio recorded and transcribed (identity-masked email feedback and interview transcripts are available at http://architecture.mit.edu/house_n/data/languagelearning/).

Both participants were given a post-study quiz on their aural comprehension of the presented vocabulary. The male participant was initially quizzed alone. After each phrase was played, he verbally provided the English translation. He could ask that a phrase be repeated up to two times. In this session, he was quizzed on 130 of the phrases. However, it was determined that the extended quizzing was logistically and mentally taxing, so the quiz was broken into two sessions. In the second session, both participants were quizzed at the same time and wrote down their responses. The male participant was quizzed on all 400 phrases. The female participant was only quizzed on the second day, on a subset (270) of the phrases.

4 Results

The participants used the application on 26 different days between October 5-November 1, 2006, for a total of 120.9 hours. The mean daily run time was 4.6 hours (SD=2.0). The longest consecutive run time was 6.67 hours. The participant muted the application 13 times on 5 evenings (for an average of 78 minutes).

The male participant reported that he turned on the phone and application as soon as he remembered after getting home from work. He initially did not run the application in the morning because he was in rush (*"I burst out of the house so quickly"*) and because he did not want to disturb his wife, who was still sleeping while he was getting ready. On weekends, the participants woke up later and were often at home for a longer period of the day. On these days, they turned on the phone when they got up and turned it off when they went to sleep.

With the settings used during the study, the application could run for about 5 hours before the phone ran low on battery power. The phone dipped to ≤10% battery charge five times, after the application was run for between 4.8 to 6.1 hours. The male participant reported that the battery power usually was sufficient and that on the few occasions when it was running low, it was easy to charge the phone wherever he was.

The male participant typically set the phone on a surface near to where he was doing an activity and left it unplugged. He identified the study area near the computer, the countertop between the kitchen and living room, and the living room couch area as common places where he would leave the phone. If the phone was left in the study area, the audio play was still loud enough to hear in most parts of the apartment.

Phrases were played 6,926 times during the study period, on average 57.3 times per hour and up to 142 times in one hour. While the application was running, activations were detected from 70 wireless object usage sensors, 80 RFID tags, 22 switch sensors, and 6 flow sensors. A subset of these activations triggered phrase play. Up to 400 different phrases were available to the application, but only 274 were actually presented. If multiple words were activated within a narrow time window, only one was randomly selected to play. Additionally, some objects in the home were not interacted with or not in a way to trigger sensor activation. Phrases for particular objects were played between 1 and 578 times, with a mean of 25.3 (SD=63.9). Frequently played phrases included toilet (578 times, see note later), computer keyboard (530), computer (447), sink (292), chair (274), and faucet (212). When asked to estimate the total number of different phrases that they heard, the male participant guessed 100 and the female participant guessed 75.

4.1 Pilot Learning Performance Results

Both participants performed better on audio comprehension of phrases that were presented during the study than on those they had not heard during the study. The 126 phrases that were not played during the study were used as "control phrases" when evaluating performance. The male participant guessed correctly on 57.7% of the presented phrases (158/274) and 51.6% of the control phrases (65/126). The female participant guessed correctly on 39.9% of the presented phrases (71/178) and 28.3% of the control phrases (26/92).

Table 1 presents pilot performance data for each participant grouped by frequency category. These data suggest that the more highly played words or phrases were learned better than phrases played less frequently. A point biserial correlation coefficient of $r=0.19$ ($p < 0.01$) was computed for the combined performance data for the participants with respect to the raw number of times each phrase was played. This suggests a weak, but statistically significant relationship between the frequency with which each phrase was played and whether the participants recalled the correct translation. It should be noted however, that both participants recognized several of the control phrases, which were not played during the study; in some cases, these phrases were cognates of phrases that were played, but in other cases, the participants likely knew the phrases beforehand.

Table 1. Pilot performance data for each participant; scores reflect percentage correct for the subset of phrases tested for different frequency of play categories

Times Played	Male's Scores		Female's Scores	
20+	54/68	79.4%	28/45	62.2%
8-19	40/63	63.5%	19/42	45.2%
4-7	30/64	46.9%	15/45	33.3%
1-3	34/79	43.0%	9/46	19.6%
0 (control)	65/126	51.6%	26/92	28.3%

4.2 Usability and Acceptability

The male participant's initial reaction to the language tool, provided via email, was surprise at its accuracy and the eerie quality of having a system sensing one's actions, *"it's almost creepy, like the phone knows what you're doing! But it seems like an awesome way to learn a language."* At first, it seemed to present phrases too frequently, *"the phone might speak a bit too often, it was almost constant (well, not quite, but it seemed it) at a few points."* After 10 days, the participant concluded *"most of the time though, it's a welcome interruption. Every now and then I wish I could bring it along with me somewhere just to hear it say different words!"*

The participants noted that the phrases seemed to get into a kind of queue at times, with a noticeable delay between an action and a stream of phrases related to that action. This resulted from the pilot application's phone network latency. They estimated that most of the time, though, the pairing was *"pretty fast,"* with a phrase playing within 10-15 seconds following an action, although it was difficult for them to judge. At times, the participants would wonder why there was a burst of phrases, corresponding with their actions, followed by a relative silence for a period of time (a likely result of the filtering heuristics, limiting repeated phrase play).

The participants were asked to describe situations when phrase play was irritating. The male participant noted that he had muted the application a few times when watching movies or on the phone. However, having the application on in the morning was the most difficult, *"I guess because I was in the shower, and trying to relax, and still waking up, the phone kind of got to me."* The female participant described a day working at home and listening to audio transcripts; the language tool, using an auditory medium, disrupted her focus on work, *"it would drive me nuts."*

However, the participants noted for the most part, the "background" quality of the application made it easier to selectively attend to or ignore. The male participant commented, *"it was nice that you can do it while you are doing other things. You can kind of choose your level of attention. If you thought you were busy or stressed, then you could kind of ignore it, but if you were relaxed, you could kind of listen to it more. So it seemed very flexible to me in that sense. Whereas I got a language CD once and played it in the car, I was driving, but I felt like I was forcing it on myself."*

The female participant noted that the idea of having an always-on application sounds like it should be burdensome, but the four-week experience convinced her otherwise, *"I think that if someone had said, outside of this experiment, 'you're going to have this phone ... and the language is going to be based on things that you move and it's going to say the words,' I would think, well, this time and this time during the day, I don't have much going on, and I can concentrate. But now that, since we've had it running in the background, I found no problem with that."*

Given that the application had this "background" quality, the participants were asked what determined their level of attention to phrase presentation. The male participant described that he would ignore the phrases, *"when I was mentally engaged, like reading or when I was writing stuff, then I would miss words."* He would attend to the phrases when he was active, but doing more physical tasks, *"I actually found, when it was in my pocket, and I was moving around doing things, I felt like my attention was pretty high then, mostly."* Both participants felt that being the subject of the phrase aided learning and comprehension, *"I definitely would hear words better if it was something that I was using, as opposed to when it was something that [my spouse] was using."* Sometimes the application was easy to ignore; when they had a visitor over, he didn't seem to notice, instead quickly settling in to watch a movie with the participants. The researchers noticed a similar phenomenon in pilot testing, when bystanders seemed oblivious to the phrase play.

The participants indicated that the application was not disruptive to their conversations. They often took a playful approach, repeating phrases they liked, such as "fregadero" (sink), and teasing each other when a phrase revealed something one person was doing, out of sight of the other. *"I would get a glass of water and [my spouse] would be like 'can I have some water too?' (laughs)."* The English phrase was presented first, making it possible for the female participant to try to *"beat the machine"* by speaking the translation before it was played.

Because the male participant was carrying the phone, the female participant typically did not have the opportunity to see the phrase or ask for it to be repeated. In the last week, she had the application running in the morning after her husband had left for work. *"That's when I realized, I'm much more of a visual learner, because if I can see the word, I can remember it, a lot better than when I just heard it."* The male participant agreed, *"it's almost like when I see the word, then it would sound different the next time. But generally, when I would see it just one time, it would kind of click."* He primarily used the "repeat" option to see the phrase, if it was unfamiliar.

4.3 Mental Model of the Sensing System

The participants reported that they rethought their understanding of the sensing system after the application was introduced. Played phrases led them to question where sensors were placed, what the sensors "knew" about their actions, and the range of data transmission.

When they installed the wireless object usage sensors at the beginning of the study, they were given instructions that described how the sensors worked and provided recommendations on how to optimize detection through careful placement. They lived with the sensors for six weeks before running the language-learning tool, but once they were able to hear words play in correspondence with their actions, they were able to investigate how sensitive the sensors actually were. For example, the male participant described experimenting with jostling the toilet paper to determine what degree of movement would trigger phrase play. *"Because it [the sensor] wasn't on something that was moving, and I didn't think it would get anything. But …it wasn't just random – and I tested it out a few times, because I kept hearing it."*

For the first few days with the language application, the participants reported that they moved more objects, such as the broom, just to discover which objects had sensors and labels. They discovered RFID tags that they didn't know were there or had forgotten about. *"I think I noticed the ones on the shelves in the kitchen earlier, but I totally forgot about them until the language program, because I think they said bowls, and we were looking at the bowls, and there's nothing on the bowls! Then we noticed the RFIDs underneath."*

The participants were not sure whether the RFID-tagged objects were also contributing to phrase play. They realized that some objects with wireless object usage sensors triggered multiple phrases, including those referencing objects inside or related to the primary object. In one case, the phrase "egg - huevos" was played when the female participant was actually getting eggs from the refrigerator. After a careful inspection, she determined it wasn't a case of a "hidden" sensor, but rather a coincidental pairing of phrase and action. They also discovered that interaction with an object might also trigger a nearby object, presumably due to bumping or vibrations. *"I would open a cabinet and it would say 'cabinet,' but it would also say 'sink' or something. It felt like that area [the kitchen island] was all interconnected."*

The participants also discovered situations where labels were not correct or sensors were not functioning. For example, the phrase for "antiseptic" was played when accessing the medicine cabinet, but no antiseptic was contained within. All chairs but one in the dining room triggered "chair - silla," due to a failed sensory battery. On the night when they left the application on for a while, they were surprised not to hear "bed" (again, likely due to a failed battery).

The language tool also gave the participants a greater awareness of the apartment and each other's routines. Typically, the participants would turn off the application just before going to bed. One night, they left it on as an experiment. They continued to hear phrases, which they presumed were being activated due to vibrations in the house. *"It was interesting, it kind of gave you a feel for how many things are in motion, realizing that the shelves are vibrating and the toilet's vibrating."*

They frequently heard a stream of translations for "cup, saucer, placemat," which helped them locate the source of a funny vibration sound in a cabinet, perhaps related to the under-cabinet lighting. The participants determined that the toilet had a leak, and was continually running, because they heard "toilet - taza del baño" more than any other phrase. The female participant described one incident where the language tool made her more aware of something she had forgotten to do. *"It was actually helpful one time when I left the burner on, and it would say burner every once and a while, so we'd gone to sit down for dinner, and I said, 'why's it doing that?' and the burner was on (laughs) so I turned it off."*

The introduction of the language tool reminded them that their activities were being recorded. *"I definitely became more conscious of being recorded again when the language phone came into play. It was making it very obvious, you lifted your glass and it said 'glass.'"* The male participant described how having the objects he was using echoed back, and knowing that his spouse was possibly hearing those phrases too, made him more self-conscious of his routine in the bathroom. *"I might have been a little bit conscious about that a few times, like making sure I'm shaking the soap so it's clear that I'm washing my hands. It definitely reminds you of things."*

The female participant reported that she became more aware of some of her husband's routines that normally she did not observe first hand. In the morning, when she is still resting, her husband is getting ready for work. *"I got more aware of [my spouse]'s morning routine, when [he] had it on in the morning. I hear the movement around in the morning, but I don't know what he's doing, but when that was on, 'oh, he did this, and then he did this.'"* She described how she assumed that when he was working in the afternoon and evening in the office, that he stayed in one place at the computer. While she was in the kitchen or living room, however, she could hear phrases that indicated he was going into the bedroom or interacting with the stereo.

The participants reported that they were comfortable with having this new awareness of their home and each other's routines, and even found humor in some of the situations that arose, but when asked to describe what it would be like to have phrases corresponding with their actions transmitted to relatives or vice-versa, they had mixed feelings. Both participants imagined that it would be fun for kids to know a little bit about what their relatives were doing and the female participant imagined that it would be a good way to informally monitor older adults or kids at home alone. *"It would be good in one way, because you would be thinking about them, like 'Why are they using the water heater again?' (laughs)... I could see it as an interesting way to keep tabs on them in a non-conscious way, learning the vocab. But then also like, 'it sounds like their stove is on…'"* The male participant, however, worried that it might lead family members to judge each other; *"I could see myself saying to my mom, 'you were reading that book for 13 hours,' and she could say 'you were on the computer for 14 hours'."* He also noted that it would remove plausible deniability with respect to not answering the phone; *'how come you didn't answer your phone the other day? I know you're at home, because you were on the computer' (laughs)"*

4.4 Ideas for Re-design

During the study, the male participant suggested that he would like to be able review phrases that had been played on a given day, *"if I'm in the middle of doing things, I don't see the screen, so it's hard to know exactly how things are said correctly without* *seeing the word in print. When possible, I press the button to have a word repeated, but sometimes I can't (when my hands are full), so I'd like to be able to review them later."* After the study, he decided that it would be something he would do once or twice a day, when he had a spare moment working at the computer or at dinnertime. The female participant suggested that, *"I feel like it would be useful, since it's a program that's on the phone, if you could take the phone with you and review the words that happened, riding [public transit]."*

Both participants liked the diversity of phrases, even though they felt there were many that they didn't hear frequently enough to learn. They did wish that the vocabulary set had gradually shifted to include verbs and phrases. They noted that some phrases, such as computer keyboard, became boring, because even though they had learned them, they were frequently repeated. The male participant suggested being able to indicate that a phrase was learned so that a new phrase (e.g., "typing") could be introduced. The old phrase could then be played again every so often to help them retain what they had learned.

As an added challenge, the female participant suggested that hearing multiple speakers say the phrase would be helpful, *"because then you could pay more attention to what the word is supposed to be."* The participants were asked if they would be willing to grade other people's pronunciations, if they could have that service done for them in return. The male participant suggested instead that the application should *"have you pronounce the words and record them, and then maybe if it mixed your pronunciations in with the other person's,"* and that would be helpful for gradually correcting pronunciation without feeling threatened.

The participants had commented that the act of taking the quiz forced them to listen more closely to the phrases than the more passive exposure they had received during the study. They were asked about a possible interactive feature, where a Spanish phrase would be played, relating to a nearby object, that they would then have to interact with or move to "guess" the answer. The male participant had some reservations about the application having an interactive component, *"What I like about it now is that it doesn't involve much from you… it is kind of background, you can carry it your pocket and set it down somewhere and still hear it. And if I had to actually like do something, because a lot of times your hands are occupied or your mind's occupied and it would be kind of disruptive to get up and shake the chair."*

However, the female participant proposed that a game-like mode would be acceptable if it could be deliberately (not automatically) initiated and in particular, if used when playing with younger members of the family. The male participant agreed, suggesting that if a prompt to initiate a game was ignorable, as it was with the mobile phone cognitive game that he tried out in the earlier part of the study, that it would be acceptable; *"like once or twice a day, I could choose the time or if it buzzed me. If it said 'alright, now it's time to play.' If I had a few minutes from what I'm doing, and then go and do it."* The female participant noted that an interactive mode would probably be more vulnerable to a novelty effect, and would be best employed when the user was actively preparing for a trip.

The participants noted that the installation of the wireless object usage sensors at the beginning of the study was a tedious task, *"that was pretty intense, all those sensors. I don't know how to make it faster, because it seemed like it was as fast as it could go, but ... It was just a lot of time to spend."* When asked what it would be like to tie the language application more closely to the sensor installation, the female participant felt that it would provide a needed motivation to complete sensor setup, *"here we were just putting it on and there wasn't an immediate … knowledge of how the [sensors] were going to be useful, but maybe if we had [something that said] 'tomorrow, if you have all these [sensors] up, you are going to start learning these words."* The male participant commented, *"I could see having the language program working and I guess placing the [sensors] more strategically or even adding RFIDs to some things… if there were things you wanted to learn, put the [sensors] right on there. And I could see doing it gradually... but it might not seem so bad - not going around the house searching for things, but as we used them."*

The participants concluded that they would like to have this kind of application in their own home. They envisioned using the application for two months prior to traveling to a foreign country, at about the time when they usually book their flights. During that two-month period, they would see themselves running the application continuously. The male participant commented, *"Even at four weeks, it wasn't getting that old. And that was, even with one set of words, so I imagine if the vocabulary evolved over time, it could stay fresh for a while."* When asked how many words they would expect to learn during that period, the male participant described spending two months in Greece and estimating afterward that he had learned about 500 words. Correspondingly, he would want the language tool to help them be ready for a vacation, with two months of prep time, by learning about 500 words, and ideally up to 800-1000 words, in order to function well as travelers in the culture.

5 Discussion

Ubiquitous sensing systems are becoming sufficiently robust and flexible to support the prototyping of context-sensitive applications. Live-in laboratories and portable research kits provide the opportunity to test out these applications in real life situations for an extended period of time. However, these studies are necessarily limited in the number of participants, and therefore results must be interpreted as exploratory. The participants in this study were optimistic about possible uses of technology, though they were not experts. By agreeing to participate in a study where their everyday routines were recorded for research, they demonstrated a comfort with ubiquitous sensing that may not be shared by most individuals.

Several design decisions were made to focus on exploring the usability and acceptability of an always-on microlearning application. The vocabulary set was limited and was presented simply, as L1-L2 pairs. Spanish was chosen as the second language despite the participants having some previous knowledge of it and its closeness, through its phoneme system, cognates, and loanwords, to English. Comparison conditions, with learning out of context, were not available. The learning performance quiz was conducted out of context and with a less formal procedure than would be used in a large-n experimental study.

Given these study attributes and design decisions, we were able to investigate how a young couple, who would be likely to consider purchasing ubiquitous technology as it becomes available, would respond to an always-on interface designed to help them with the task of learning vocabulary in preparation for travel. Because they were able to live with the application for a one-month period, having it run for over 120 hours and hearing almost 7000 phrases, they experienced a variety of situations where the interface was either compelling or irritating, accurate or unexpected. The application became part of their everyday life and featured in their conversations and activities.

Other protocols were run simultaneously during this study, which inevitably affected the participants' evaluation of the prototype. For example, the participants took it as a given that their activities were being sensed. However, they also had a more comprehensive experience of the system. Having placed and labeled many sensors themselves, and with a greater understanding of what exactly was being detected in their daily routines, they were able to weigh the benefits of the application with the burdens of setup and being monitored. Having tried out mobile applications earlier in the study, they were able to compare the "background" aspect of the application with more interactive interfaces. After four weeks, the participants had more expertise on the use of the system than the researchers and were able to provide articulate feedback and ideas for extensions.

The learning performance gains evidenced by the quiz suggest what we would expect – that more exposure to words would lead to more learning. They also show, however, that a phone-based interface may be able to support learning, even for the difficult task of aural comprehension of foreign language vocabulary. However, the gains, especially when factoring in the participants' previous knowledge of Spanish, were relatively small, and certainly below the participant's learning expectation of 250 words per month. This may be attributable to both the simplistic heuristics used to determine phrase presentation and the passive quality of the interactions.

Several techniques may make phrase presentation more effective. Using the principle of spaced repetition, phrases could be played very frequently at first, at every interaction with an object, and then increasingly less often as the phrase becomes familiar. As the participants suggested, aural presentation could be reinforced with more opportunities for reviewing information visually. This could occur in periodic prompted review sessions, such as while riding the bus; through subtle placement in other media, such as having presented words automatically translated on viewed web pages; or with novel display techniques [21]. The elimination of the latency introduced by the current phone implementation might also allow a higher density of words to be presented at the same level of user burden, potentially increasing learning effectiveness. Moreover, near zero latency creates new interface opportunities, such as double tapping an object to hear a word again.

Although the non-demanding background quality of the interface was viewed as a benefit, adding more interactive features, such as giving the user the ability to review phrases or to specify when a phrase has been learned, as well as periodic quizzes or games, might aid learning. Applications that are delivered on the PC or phone typically receive explicit user input and usage pattern data that can help determine content and interaction pacing to match the learner's needs. A ubiquitous interface such as the one described in this work must find new ways to obtain this information without tipping the scale back to burdening the user with forced interactions.

Applications that employ immediate context-sensitive feedback may provide a motivation to purchase, train, and maintain ubiquitous sensing systems, even those that use dense-object sensing consisting of hundreds of miniature sensors placed on objects throughout a home. The participants demonstrated more insight and interest in the working of the sensing system after the language tool was introduced, and through playful interaction they were learning about the strengths and limitations of the sensor system, as well as information that would be helpful to maintain it over time (e.g., strategies to identify if and when something is not working). Future work needs to explore how the immediate rewards of an application such as this can be leveraged to help enable installation, use, and maintenance of other sensor-enabled applications with delayed return-on-investment, such as longitudinal personal health monitoring.

The observations from this study suggest that, at least for some people, it may be possible to layer context-sensitive interactions onto everyday routines without them being perceived as burdensome. Although it seems like the high density of messages (up to 142 audio-clips in one hour) would be disruptive, the participants in this study found it possible, for the most part, to choose their level of attention to the always-on interactions. It may be that the repeated exposure, which appeared to be associated with learning rates, also lessened the demand quality of the interactions. The congruence between the message and the participants' immediate context may have also reduced the perceived disruption. The participants liked the "background" quality of the interface and considered what they were doing "learning a language." This suggests that the non-demanding aspects of always-on microlearning applications may be valuable for introducing users to involved tasks, without risking burnout.

We have used the interview feedback from this study to consider design attributes that may be applicable to other context-sensitive, always-on systems that need to sustain user interest, while avoiding unnecessary disruption. These attributes include "background" audio presentation with optional, visual feedback; preferential presentation when users are engaged in physical tasks or are moving through a space; repetition and context congruence to support selective attention; and built-in opportunities for review, quizzing, and signaling readiness for new information, perhaps in idle moments, such as during a commute. Always-on systems that encourage the users to experiment, to care about correct mappings, and to get new content in exchange for extending or elaborating a ubiquitous system may be easier to maintain. Finally, systems that provide simple and playful interactions, encourage self-awareness of personal routines, and give users extended, non-demanding exposure to a more complicated task may be perceived as having immediate value and may justify the introduction of novel ubiquitous computing in the home.

Acknowledgments. This work was supported by Intel Corporation. We thank our participants for the generous contribution of their time and feedback, and TIAX LLC for collaboration on the PlaceLab. The sensors used were developed with support from National Science Foundation grant #0313065. The PlaceLab stay was funded by Microsoft Research.

References

1. Hug, T., G. Gassler, and C. Glahn, Integrated micro learning - an outline of the basic method and first results, in Proceedings of Interactive Computer Aided Learning. 2004: Kassel Univeristy Press. p. 1-7.

2. Hug, T., Micro Learning and narration, in Fourth Media and Transition Conference. 2005: Cambridge, MA.
3. Dempster, F.N., Effects of variable encoding and spaced presentations on vocabulary learning. Journal of Educational Psychology, 1987. 79(2): p. 162-170.
4. Tulving, E. and D.M. Thompson, Encoding specificity and retrieval processes in episodic memory. Psychological Review, 1973. 80(5): p. 352-73.
5. Davies, G. and D.M. Thomson, Memory in context : Context in memory. 1988, Chichester [England] ; New York: J. Wiley. ix, 359 p.
6. Godden, D.R. and A.D. Baddeley, Context-dependent memory in two natural environments: On land and underwater. British Journal of Psychology, 1975. 66(3): p. 325-31.
7. Smith, S.M. and E. Vela, Environmental context-dependent memory: A review and meta-analysis. Psychonomic Bulletin & Review, 2001. 8: p. 203-220.
8. Brown, H.D., Principles of Language Learning and Teaching. 2nd ed. 1987, Englewood Cliffs, N.J.: Prentice-Hall. xvi, 285 p.
9. Importance of learning a second language survey, in Gallup Poll. 2001.
10. McGinnis, S., The less common alternative: A report from the task force for teacher training for the less commonly taught languages. ADFL Bulletin, 1994. 25(2): p. 17-22.
11. Nation, I.S.P., How large a vocabulary is needed for reading and listening? The Canadian Modern Language Review, 2006. 63(1): p. 59-82.
12. Faaborg, A. and J. Espinosa, Using common sense reasoning to enhance language translation with mobile devices. Last accessed March, 2007. http://agents.media.mit.edu/projects/globuddy2/
13. Thorton, P. and C. Houser, Using mobile phones in English education in Japan. Journal of Computer Assisted Learning, 2005. 21(3): p. 217-228.
14. Wilson, D.H. and C. Atkeson, Simultaneous tracking & activity recognition (STAR) using many anonymous, binary sensors, in Proceedings of PERVASIVE 2005. 2005, Springer-Verlag: Berlin Heidelberg. p. 62-79.
15. Philipose, M., J.R. Smith, B. Jiang, A. Mamishev, S. Roy, and K. Sundara-Rajan, Battery-free wireless identification and sensing. IEEE Pervasive Computing, 2005. 4(1): p. 37-45.
16. Munguia Tapia, E., S.S. Intille, and K. Larson, Activity recognition in the home setting using simple and ubiquitous sensors, in Proceedings of PERVASIVE 2004, A. Ferscha and F. Mattern, Editors. 2004, Springer-Verlag: Berlin. p. 158-175.
17. Intille, S.S., K. Larson, E. Munguia Tapia, J. Beaudin, P. Kaushik, J. Nawyn, and R. Rockinson, Using a live-in laboratory for ubiquitous computing research, in Proceedings of PERVASIVE 2006, K.P. Fishkin, et al., Editors. 2006, Springer-Verlag: Berlin Heidelberg. p. 349-365.
18. Fishkin, K.P. and M. Philipose, Hands-on RFID:Wireless wearables for detecting use of objects, in Ninth IEEE International Symposium on Wearable Computers (ISWC 2005). 2005. p. 38-43.
19. Patel, S.N., J.A. Kientz, G.R. Hayes, S. Bhat, and G.D. Abowd, Farther than you may think: An empirical investigation of the proximity of users to their mobile phones, in Proceedings of UbiComp 2006: Ubiquitous Computing, P. Dourish and A. Friday, Editors. 2006. Springer-Verlag: Berlin Heidelberg. p. 123-140.
20. Beaudin, J.S., S.S. Intille, and M. Morris, MicroLearning on a mobile device, in Proceedings of UbiComp 2006 Extended Abstracts (Demo Program). 2006.
21. Intille, S.S., V. Lee, and C. Pinhanez, Ubiquitous computing in the living room: Concept sketches and an implementation of a persistent user interface, in Proceedings of UBICOMP 2003 Video Program. 2003.