CS6750 – Assignment M5

Sergiy Palguyev

spalguyev3@gatech.edu

Abstract—For multitasking individuals that engage in physical activities while listening to MOOC lectures the iPhone Udacity application demands constant, physical user interaction. To bridge this Gulf of Execution while maximizing user safety, attention and information retention, this project shall explore supplemental interface features for effective studying while multitasking.

1 QUALITATIVE EVALUATION

The qualitative evaluation plan consisted of a survey evaluation, due to the textual prototype presented in Assignment M3. The prototype of interest was the NLP enabled software solution presented in M3 as a textual prototype. This prototype is a software piece focused on Natural Language Processing. This application has no advanced Artificial Intelligence software. Only by learning and analyzing the voice commands from the user, the system adapts to better capture each command. Otherwise, the application simply exposes all possible Udacity functions such as navigation, playback controls, and video capture controls. Additionally, the application allows for interfacing with 3rd party applications like Piazza or notes taking in order to provide the user all the tools necessary for productive learning environment without having to switch context of usability or control.

Thirty-five participants answered the survey. All participants could take the survey asynchronously. All surveys were answered in a timely manner but certain open-ended questions led to reevaluations and further concept refinements, previously unforeseen. For further research, interviews and detailed wireframes would be created in order to create more appropriate visual aids for survey takers to assess. Additionally, interviews will be implemented as part of the analysis in tandem with the surveys in order to provide a more intimate evaluation of the prototype from the user's personal perspective.

Report

The following statistics are summarized from the survey done by OMSCS students and other participants throughout the past seven days. Over 60% of the responders responded to exercising 3-5 times per week while listening to audio material at the same time. More than 2/3 of the responders felt satisfied with touch screens as the main interface with their phone while 1/3 felt either neutral or dissatisfied with the interface. Almost 80% of responders multitask between 2-4 tasks on their phone at any one time. A multiple choice question on limiting factors for multitasking, led to the fact that participants are limited by 3 general categories: the physical dimensions of a small screens, the interactions of applications with either the operating system or with each other (considering multiple applications are used), and the human limitations of cognitive load activities which always bade for attention. Lecture material is equally well absorbed visually as well as auditory by participants and most participants spend the late hours of the day on lecture material. The majority of participants would also prefer to exercise or eat in tandem to watching lectures. The raw results are attached in Appendix A.

Analysis

The takeaway from this survey has provided some very important points considering my prototype and possible future design improvements. Most participants multitask and do so in the late hours while either eating or exercising. Application constraints are the biggest factor limiting the participants from successful multitasking, where the applications either have poor interaction between each other or the operating system, forcing the user to lose focus from absorbing material. One of the most surprising results from the survey was exactly how many people do multitask, and more specifically, the amount of tasks they context switch between. This fact brings to light the assumption made the abstract, that an extension of smart phone technology is required in order to facilitate better multitasking, but more importantly that the technology is currently absent while the market demand is quite high.

As such, my prototype lifts certain constraints present in applications as the software solution allows voice-command interaction, freeing the physical limitations of the user and implements notes and application context switching in order to facilitate smoother multitasking techniques. Some improvements to the prototype obviously present themselves form open ended questions of the us-

ers. For example, under-the-hood features, such as short term memory of where each application stopped its place (i.e. video-pause, etc.), speed of context switching, traversal not only of the application, but between applications, and cognitive load on remembering the necessary commands are all considerations for improvement. Making commands more akin to natural language, storing certain moments of an application in memory so an app does not reset after context switching, and improving speed of transitions are all possible design features to be implemented in the future.

2 PREDICTIVE EVALUATION

The predictive evaluation is conducted on the same NLP enabled software piece as the Qualitative Evaluation. The user would have three goals to complete involving two context switches and a list of voice commands that the software will recognize. The goal of the evaluation is to evaluate the usability of NLP software for conducting established tasks, such as navigating Udacity, as well as performing newly enabled tasks such as note-taking during a video.

The user is required to interact with their device physically only once, opening the Udacity application. This is required in order for the NLP software to enable and begin accepting commands. In order to maintain privacy, the NLP software will only listen to the user commands when addressed as "Buzz". In order to speed up navigation, "Buzz" will listen to the user for 5 seconds to allow navigating through multiple menus without saying "Buzz" in front of each command.

Goal: Watch a video

1. The user knows that Udacity requires a login and thus the natural commands to follow enter the Udacity page would be "Buzz, enter Email address: Bob@email.com, Password: 12345 and Login".
   The software will recognize the words "Username", "Password", and "Login" as separate keywords with the appropriate input, proceeding to Log in to the website.
2. Seeing the Log In page, the user would say "Select My Classroom" in order to enter the "My Classroom" page of Udacity.
3. Through proper mapping and simple keywords of the NLP application, the UI of the website itself lend to the most obvious keywords a user

would state. As such, the next step to continue, the user would say "Continue Human-Computer Interaction Course".
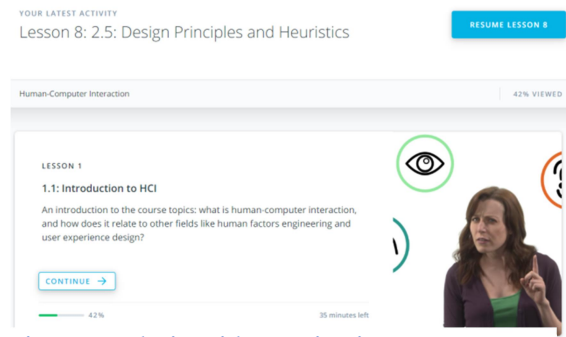

Figure 1: Udacity video navigation

4. At this point, the user can have multiple deviations in navigation.

   Initially, the user may desire to continue where previously paused, and that would simply be accomplished with the command "Resume Lesson".

   Alternatively, the user may desire to choose another lesson since the flow of this specific class is not synchronous with the lessons. As such, the user may say "Play Lesson 3.4" or alternatively the user may ask to "Scroll Down". If deciding to "Scroll Down", the software would scroll down one page, and wait for the user to ask to "Scroll Down" again. When the desired lesson is reached, then the user again may say "Play Lesson 3.4".

Goal: Record a note while watching a lecture

1. While the video is playing, a user may want to record notes for themselves. At this stage, the user may want to continue a previous note or start a new one, and "Buzz" would accept both commands. The corresponding commands would be "Buzz, continue previous note" or "Buzz, take new note".

2. A small notepad would pop-up with a screenshot, title, and timestamp of the currently running video.

3. Any words dictated at this time, would be transliterated on to the notepad.

4. The commands "Buzz, resume video" or "Buzz, close notepad". Would save the note and resume the video lecture two seconds prior to the note being taken.

    Saying "Buzz" in front of the command is necessary for the software to differentiate the text being dictated from the commands to "Buzz".

    In addition, starting the video form 2 seconds prior to the note would allow the user time to regain conscious placement in the video lecture prior to interruption.

Goal: Post or search piazza while watching a lecture

1. While the video is playing, a user may want to post a new question or search the content of Piazza for clarification on a topic or concept discussed in the video. The user would then say "Buzz, open Piazza".

2. The user knows that Piazza requires a login and thus the natural commands to follow would be "Enter email address: Bob@email.com, Password: 12345 and Login".

3. Now the user has two choices, either post a new post or search for posts containing word of interest. For a new post, the user would say "New Post to Entire Class" or alternatively "New Post to Instructors". This would open a new post page with the appropriate target audience selected.

4. The user would then say "Summary: This is a summary". Following by "Details: This is text about the details".

5. The user would then say "Post my question." In order for Piazza to post the user question to the Forum.

6. If, on the other hand the user would like to search the forum, they could say "Search for keywords M2 Assignment". The NLP software would then interpret the dialog after the word "keyword" as the keywords to search for, providing the search results to the user.

7. The command "Buzz, resume video" or "Buzz, close Piazza" would resume the video lecture two seconds prior to the note being taken.

    Saying "Buzz" in front of the command is necessary for the software to differentiate the text being dictated from the commands to "Buzz".

    In addition, starting the video form 2 seconds prior to the note would allow the user time to regain conscious placement in the video lecture prior to interruption.

Voice control allows not only for complete Udacity website manipulation, but additionally manipulation of third party applications to enable notes and Piazza interactions thus greatly improving simplicity of operation while maintaining proper mapping of functions of each application. All functionality is relegated to voice commands which do not require any visuo-spacial interaction from the user, unless desired. This prototype maintains a low Gulf of Execution and Evaluation. Each command has an expected and direct result. Since there is no artificial decision making implemented in the application, the user expects commands to execute exactly, and can determine failed commands or misinterpreted commands immediately for instantaneous feedback. Structure, Feedback and Consistency are all maintained as mapping of functions to voice commands stays rigid without AI reinterpretation, but rather through training of natural human speech. Additionally, Easy of Use and Comfort are maximized since users can maintain and expand their multitasking lifestyle while maximizing the information retention and usability of the lecture material.

3 EVALUATION SUMMARY

For the next iteration of the design lifecycle, the most pressing question would be whether the current prototype is enough to improve the quality of interaction between the user and the task. It is feasible that the prototype may be good, but whether the improvement is worth the effort and investment is not clearly understood. If the improvement in quality is marginal, the project may not be worth the investment. Would a combination of 2 or more prototypes be a better solution than just one or the other? Is there user fatigue associated with one type of interface, even if it may be initially simple to use? Instead of abstracting away certain features of the base application and wrapping them in other features, would it make more sense to redesign the original application? These are all questions which could be further granulated and explored with continued design cycles.

This experience brought to mind the idea of user fatigue. It is feasible that a Voice Command interface may be the simplest form of an interface. But not all command structures are natural to all people. For this reason an initial user evaluation may be positive, but after using such an interface for hours, or days, the interaction may become tedious. A physical interaction, on the other hand, maybe force the user to be more physically involved, but its simplicity may

prove to be more beneficial to the user. As such, an idea for a hybrid prototype has emerged. A hardware controller with a minimalistic design and minimal hardware buttons may be designed together with a microphone and a built in NLP software package which could also accept all command directly as voice commands. With this sort of interface, the user would have flexibility of control and could switch back and forth at their own desire. This would, additionally, provide a quicker time to international markets, as people who do not know English voice command would still be able to user the controller for all their navigation and multitasking needs.

The biggest revision to the NLP prototype would be to incorporate an expanding data storage and retrieval database as well as a sophisticated machine learning algorithm which would allow different human speech phrases to be associated with actions to allow a more natural language flow. This would prevent users being forced to remember commands, and instead interact with the interface in a natural speech pattern.

Assuming the changes are made, an empirical evaluation could be conducted, with many complex goals set for the users in order to assess ease of use, successful completion paths and problem areas which would need to be worked on for future prototypes. Initially, interviews with frequent users would gain insight into prototype feasibility, with finally the empirical evaluation providing concrete evidence of prototype viability.

4 APPENDIX A

1. How often do you perform a physical activity while also listening to audio material (music, audiobooks, lectures, etc.)?
2. How satisfied are you with "touch" being the main interface on your phone?
3. In general, how many tasks do you switch between when using your phone?
4. In general, how many different activities can you multi-task at any one time?
5. What, if any, technological features or constraints limit your ability to multi-task?
6. How do you best absorb lecture material?

7. At what time during the day do you watch lecture materials?
8. In general, which activities do you feel comfortable in combining for multi-tasking?
9. Which activity would you want to perform while also absorbing lecture material?
10. The current implementation of MOOC platforms is adequate teaching complex concepts. (1-Very Dissatisfied, 5 – Very Satisfied)

response	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	3-5 times/week	Neutral (No hard opinion on touch screens)	2-4	2-4	not sure. cant think of any	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation);Live lecture with drawings and visual cues	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Watching;Listening and Watching;Exercising and Watching;Driving and Listening;Occupational Work and Listening	Exercise;Eat	5
2	0-2 times/week	Satisfied (Do not know what people are complaining about)	2-4	5+	small screen size	Visual (Watching the lecturer, including hand gestuer and visual presentation);Other	Early Morning (6am-9am);Mid-Day (10am-1pm);Afternoon(1pm-4pm);Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Watching;Listening and Watching;Exercising and Listening;Exercising and Watching;Driving and Listening;Occupational Work and Listening;Other	Exercise;Work;Eat;Sleep	4
3	3-5 times/week	Neutral (No hard opinion on touch screens)	2-4	Only 1	None. I'll perform one task across multiple applications but try not to multitask.	Live lecture with drawings and visual cues	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Watching;Exercising and Listening;Driving and Listening	Exercise;Eat	3
4	0-2 times/week	Neutral (No hard opinion on touch screens)	2-4	2-4	I only have 2 hands - But since you said technological, I suppose that means my eyes and voice are not	Other	Early Morning (6am-9am);Mid-Day (10am-1pm);Afternoon(1pm-4pm);Dinner-time (5pm-8pm);Late at	Eating and Reading;Eating and Watching	Exercise;Take care of kids;Eat;Drive;Sleep	4

#	Frequency	Satisfaction			Concern	Format	Time	Multitasking activities	Activity	Rating
					sufficient to fully support multitasking		Night (9pm-12am)			
5	3-5 times/week	Very satisfied (Sometimes I wish my significant other was made of touch screens)	2-4	2-4	none	Live lecture with drawings and visual cues	Dinner-time (5pm-8pm)	Eating and Reading;Eating and Watching;Exercising and Watching;Occupational Work and Listening	Work;Eat	4
6	3-5 times/week	Neutral (No hard opinion on touch screens)	Only 1	Only 1	running	Audio (like an audiobook or listening to video lectures)	Dinner-time (5pm-8pm)	Exercising and Listening;Exercising and Watching;Driving and Listening;Driving and Watching	Sleep	5
7	0-2 times/week	Satisfied (Do not know what people are complaining about)	2-4	Only 1	The app reseting when it loses focus	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation);Other	Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Listening and Watching;Exercising and Watching;Driving and Listening;Driving and Watching	Exercise;Eat;Drive;Sleep	4
8	3-5 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	cognitive load	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Listening and Watching	Exercise	3
9	3-5 times/week	Very satisfied (Sometimes I wish my significant other was made of touch screens)	5+	2-4	Youtube video cant run on background	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Early Morning (6am-9am);Dinner-time (5pm-8pm)	Eating and Watching;Listening and Watching;Exercising and Listening;Driving and Listening	Other	4

10	3-5 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	Screen real-estate. Ease/speed of typing on screen vs keyboard.	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation);Other	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Listening and Reading;Listening and Watching;Exercising and Listening;Exercising and Watching	Exercise;Eat	4
11	3-5 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	-	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Exercising and Listening;Driving and Listening	Take care of kids	4
12	20+ times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	No technological features	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm)	Eating and Watching;Exercising and Listening;Exercising and Watching;Driving and Listening;Occupational Work and Listening	Exercise	4
13	0-2 times/week	Neutral (No hard opinion on touch screens)	2-4	2-4	none	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Afternoon(1pm-4pm)	Eating and Reading;Eating and Watching;Listening and Reading;Listening and Watching;Driving and Watching;Occupational Work and Listening	Exercise;Work	4
14	0-2 times/week	Satisfied (Do not know what people are complaining about)	5+	5+	The speed of switching, would be nice if it was quicker.	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Listening and Watching;Exercising and Listening;Exercising and Watching;Driving and Listening;Occupational Work and Listening	Exercise;Work;Eat;Drive;Sleep	3

15	3-5 times/week	Satisfied (Do not know what people are complaining about)	Only 1	Only 1	More of an attention thing, not technology	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm)	Other	Other	3
16	3-5 times/week	Very satisfied (Sometimes I wish my significant other was made of touch screens)	2-4	2-4	The difficulty of the task	Live lecture with drawings and visual cues	Afternoon(1pm-4pm);Late at Night (9pm-12am)	Eating and Watching;Exercising and Listening;Driving and Listening;Occupational Work and Listening	Sleep	5
17	3-5 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	children	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Exercising and Listening;Driving and Listening	Exercise;Eat;Drive	4
18	3-5 times/week	Neutral (No hard opinion on touch screens)	2-4	2-4	Going back and forth among different apps in mobile	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Mid-Day (10am-1pm);Afternoon(1pm-4pm);Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Listening and Reading;Listening and Watching;Exercising and Listening;Driving and Listening	Exercise;Work;Eat;Drive	2
19	0-2 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	no	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Mid-Day (10am-1pm);Afternoon(1pm-4pm)	Listening and Watching;Exercising and Listening;Driving and Listening	Other	3
20	3-5 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	no extra fingers	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Early Morning (6am-9am);Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Watching;Exercising and Listening;Exercising and Watching;Driving and Listening;Occupational Work and Listening	Eat	3

#	Frequency	Satisfaction			Issue	Learning Preference	Time of Day	Activity Combinations	Activities	Rating
21	6-9 times/week	Dissatisfied (Have a smartphone but would love some tactile accessories)	2-4	2-4	Navigating lectures/quizzes while working out and/or driving.	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Afternoon(1pm-4pm);Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Watching;Listening and Watching;Exercising and Watching;Driving and Listening	Exercise;Eat;Drive	2
22	3-5 times/week	Dissatisfied (Have a smartphone but would love some tactile accessories)	Only 1	2-4	n/a	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Mid-Day (10am-1pm)	Listening and Reading	Work;Take care of kids	4
23	0-2 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	No convenient mechanisms to switch or combine views within the application/interface	Visual (Watching the lecturer, including hand gestuer and visual presentation);Live lecture with drawings and visual cues	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Watching;Exercising and Listening;Driving and Listening;Occupational Work and Listening	Eat;Drive	4
24	10-20 times/week	Neutral (No hard opinion on touch screens)	2-4	2-4	Not sure	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Watching;Listening and Watching	Drive;Sleep	3
25	3-5 times/week	Dissatisfied (Have a smartphone but would love some tactile accessories)	Only 1	Only 1	none	Live oratory lecture	Mid-Day (10am-1pm)	Eating and Watching	Drive	3
26	0-2 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	The screen is not big enough to show many tasks at the same time.	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm)	Eating and Watching;Listening and Watching;Driving and Listening	Exercise;Drive	4
27	0-2 times/week	Very satisfied (Sometimes I wish my significant other was made of touch screens)	2-4	2-4	Slow task switching	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Exercising and Listening	Exercise;Eat	1

#										
28	0-2 times/week	Satisfied (Do not know what people are complaining about)	5+	2-4	Screen size	Live lecture with drawings and visual cues	Afternoon(1pm-4pm);Dinner-time (5pm-8pm)	Eating and Reading;Eating and Watching;Listening and Reading;Listening and Watching	Sleep	3
29	0-2 times/week	Neutral (No hard opinion on touch screens)	2-4	2-4	none	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Afternoon(1pm-4pm);Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Listening and Watching;Exercising and Listening;Driving and Listening;Occupational Work and Listening	Exercise;Eat	3
30	0-2 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	RAM	Other	Late at Night (9pm-12am)	Listening and Reading;Listening and Watching	Exercise;Eat	2
31	0-2 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	lack of haptic feedback	Audio (like an audiobook or listening to video lectures);Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Listening and Reading;Listening and Watching	Eat	4
32	3-5 times/week	Satisfied (Do not know what people are complaining about)	2-4	2-4	screen space and can't have different audio channels (physically - meaning I cant play two videos and understand both)	Visual (Watching the lecturer, including hand gestuer and visual presentation);Live lecture with drawings and visual cues	Early Morning (6am-9am)	Eating and Watching;Listening and Watching;Exercising and Listening;Driving and Listening	Sleep	2
33	0-2 times/week	Neutral (No hard opinion on touch screens)	Only 1	2-4	screen size	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Late at Night (9pm-12am)	Eating and Reading;Eating and Watching;Listening and Reading;Listening and Watching	Eat	2
34	3-5 times/week	Neutral (No hard opinion on touch screens)	2-4	2-4	N/A	Visual (Watching the lecturer, including hand gestuer and visual presentation)	Dinner-time (5pm-8pm);Late at Night (9pm-12am)	Listening and Watching	Exercise	4