# Illuminac: Simultaneous Naming and Configuration for Workspace Lighting Control

**Ana Ramírez Chang**
Berkeley Institute of Design and Computer Science Division
University of California, Berkeley
anar@cs.berkeley.edu

## ABSTRACT
We explore natural and calm interfaces for configuring ubiquitous computing environments. A natural interface should enable the user to *name* a desired *configuration* and have the system enact that configuration. Users should be able to use familiar names for configurations without learning, which implies the mapping from names to configurations is many-to-one. Instead of users learning the environment's command language, the system simultaneously learns common configurations and infers the keywords that are most salient to them. We call this the SNAC problem (Simultaneous Naming and Configuration). As a case study, we contrast speech and GUI interfaces for workspace lighting control on a large array of individually-controllable lights. In our design process we have used a lo-fidelity text-based study, a speech-based training data collection study, an evaluation of the deployed live system and are conducting a longer evaluation of the deployed system.

### Author Keywords
Natural Speech Interfaces, Non-negative Matrix Factorization, Environment Control

### ACM Classification Keywords
H.5.2 User Interfaces:{Voice I/O, Natural language}

## INTRODUCTION
The number of electronic devices in our environment is ever increasing. While this brings greater flexibility and control, configuring each individual device becomes ever more tedious. For example, to prepare a workplace for a presentation, one might want to close the blinds, dim the lights near the projections screen, lower the projection screen and turn the projector on. Then to prepare the environmental state for a meeting, one might turn the intensity of the lights up, open the shades, ensure the projection screen is up to use the white board and ensure the projector is off. Controlling all of these devices—the lights, the projector, the projection screen, and the blinds—to achieve a desired environmental state is quite tedious. It is widespread best practice to use activity-specific "configurations" (or scenes in the case of lights) of many devices rather than setting each device individually. The user can then invoke the configuration with a single action — a keypress or in our case a speech command.

As with any interface, an interface for controlling the environmental state in the workplace should match the user's mental model. That is, the user should only need to specify an intuitive *name* of the environmental state rather than the *configuration* of each individual device needed to achieve the desired state. In the workplace, the user should be able to say "presentation lights please" or "I'd like lights for a talk now," or any similar variation, and have the system give a similar response. They should also be able to say "meeting lights" or "whiteboard lights" and get a different response. These terms are widely shared by people, and their repeated use during training allows a system to learn them as well.

Note that this problem is more challenging than simply memorizing command strings and the appropriate device settings. In the latter case, the system will be extremely brittle, and will respond only when exact training strings are provided. By *simultaneously* learning commands and device settings, the system becomes both more robust and better able to generalize. For instance, there will be many training strings for presentations that include the word "presentation" and many other filler words, but which all specify a similar light and window shade pattern. Since the system looks for common patterns in names and configurations, the word "presentation" will be strongly present in a pattern that includes presentation light settings. Thus it is able to infer the "presentation" is a salient keyword for presentation lighting vs. "please" or "now" that may also occur in command strings. Similarly, "talk" and "presentation" will typically occur with similar patterns of lights, and the system will be able to infer that they are aliases in this context. We called this the SNAC problem (Simultaneous Naming And Configuration).

The SNAC problem also arises in home environment control (configurations for movie watching, dinner, music listening, slide shows, games, napping etc.). In our experiment below, configurations are settings of groups of lights. But in more general environments, settings could include connections, e.g. a connection from X to Y could be part of a configuration called "music listening".

In addition to matching the user's mental model, we want the

Submitted to UbiWORK @ Ubicomp 2008

interface to be "calm." That is, the interface in a ubicomp environment, like environment control, should be almost invisible except during direct (focal) interaction (as advocated by Weiser [4]). In this context, a speech-based interface seems like a good option. With distributed microphone technology, the physical interface all but disappears, but jumps fluidly to the foreground when the system responds to spoken input. Furthermore, speech is often considered the most natural form of human expression and has the potential to address certain accessibility concerns.

In this paper we describe the iterative design process we have followed in designing a natural speech interface in an open-plan workspace. The system, which we call Illuminac, runs live and controls 79 devices (individually controllable lights) from 25 users who have both their own and shared environment names and configurations.

## OVERVIEW OF WORKSPACE LIGHTING CONTROL

Many open-plan workspaces have large banks of lights controlled by a single light switch. This is particularly wasteful in the evenings when few employees are around, and discourages use of available daylight in the daytime. Many lights are turned on for just a few occupants, and lights next to a window cannot be turned off without turning off the lights away from the window. More granular control over the lighting in large workspaces could enable a reduction in energy consumption by allowing unnecessary lights to be turned off. Automation is sometimes an option (turning lights on and off using motion sensors [3] or turning lights off based on luminance sensors such as in daylighting systems [5]) but has several drawbacks: lack of level control for different tasks (reading vs. using a computer), cost and complexity of wiring and configuring sensors, inflexibility to changes in furniture position (a big problem in our lab which has a very dynamic open space), and extreme annoyance to users if lights are turned out while they are still there. We feel that users are willing and able to control their lights as long as there is a highly-usable interface available.

Proprietary granular lighting control systems for both retrofit and new construction have been available for some years [1]. Most recently, a low-cost industry standard has emerged: DALI: Digital Addressable Lighting Interface [2]– is a bus system for individual light control which has been adopted by most lighting manufacturers, and deployed in major installations (e.g. Heathrow terminal 5). It is an increasingly popular option in energy-conscious building design. However, the increase in flexibility implies an increase in control complexity. The current state-of-the-art in complex lighting control is panels of wall buttons with (often cryptic) scene names, or touch screens with complex menu hierarchies. Neither of these approaches address the issue of lighting in reconfigurable workspaces, where many more configurations are possible. There is clearly a need for more flexible, intuitive and scalable lighting control.

We are exploring speech-based interfaces for lighting control in shared workspaces, which are natural and calm in terms of user attention and effort. To use the system, users should be able to customize the system to work with names of lighting scenes that are natural to them. For example, based on our experience with Illuminac, we found that one user says "turn on my lights" to turn on the two lights over her desk, while another user says "all on" to turn on the four lights around her desk. Thus the system accepts both anonymous and personalized commands. Personalization is possible using any mechanism (e.g. a personal microphone, speaker ID) that provides speaker identification, or by the user prefacing a spoken command with their name.

To add a command to Illuminac, users train the system by first recording their command. Then, the user demonstrates the desired lighting configuration and identifies herself. The novel aspect of our system is that rather than simply storing this mapping from command to configuration, we combine the recorded speech command and lighting configuration into a common representation to provide as input for a standard machine learning algorithm. Intuitively, the system uses the learning algorithm to identify structure across the space of command-configuration pairs, not just the space of commands. Once the user has trained the system on a few examples, the user can say her command into any of the microphones in the room, and the system changes the lighting scene by applying the trained model to the user's command. Because the model is trained on commands and configurations specific to the workspace, we expect to be able to perform reasonable lighting actions with less command training. For example, when a visitor who has never provided training input to the system comes into the lab, she can try her command and potentially get reasonable behavior because regular users may have already trained the system on similar commands. Of course, if the resulting behavior is undesired, she can manually change the lighting scene, thereby giving the system another training data point suited to her.

## ILLUMINAC

### Workspace Details

Illuminac was designed for and is deployed in a 2,300 square foot open-plan shared workspace with about twenty-five regular occupants (nineteen of whom have permanent desks in the workspace; the rest have permanent desks in the adjacent room). The workspace has six graded-awareness cubicles (i.e., cubicles with walls of varying heights from full height to desk height) that occupy half the room. The other half is a multi-use space for meetings, presentations, ad-hoc team meetings or individual work.

The multi-use space has a presentation screen, a "soft space" with a couch and chairs, and a set of four computers for visitors of the lab to use. There is also a tool shop in one corner of the room. Figure 1 shows a picture and the floor plan of the room. The room has 79 individually-controllable compact fluorescent lights mounted overhead. The intensity of each of the lights can be controlled over the network via a web interface. All occupants of the lab have access to the web interface (Figure 2). They can access it from their personal computers or from one of the public machines.

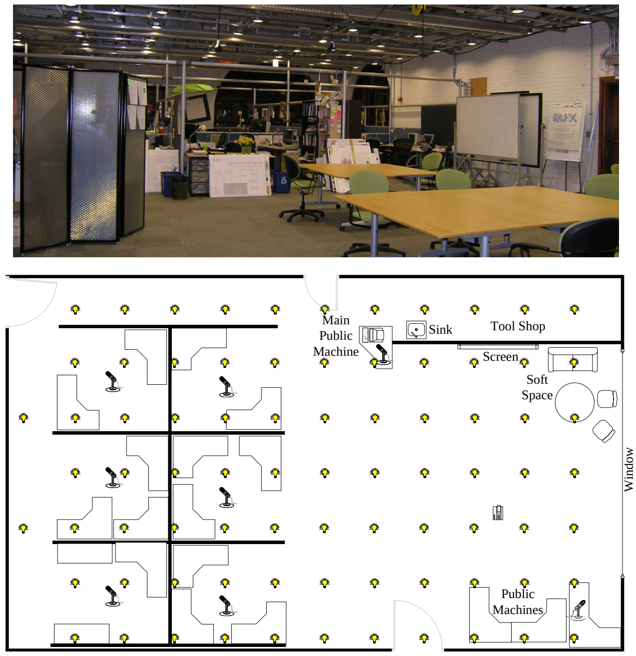


**Figure 1. Picture and floor plan of the workspace for which we have implemented a natural speech interface for the lighting control. Each of the 79 individually-controllable lights as well as the eight microphones are shown.**

The public machine next to the entrance (labeled "Main Public Machine" in Figure 1) always has the web interface open. There are eight desk microphones throughout the room to allow easier access to the lighting control system. There is one microphone in each cubicle, one at the desk in the corner, and one at the main public machine where the graphical interface to the lights is always open in a browser window. Each microphone has a clearly labeled on/off switch so that residents may control what is and is not recorded.

### Iterative Design Steps

As described in the workshop proposal, the design of ubicomp systems for the workplace requires integrating devices, systems and rules of practice. In the design of our natural speech lighting control system we used a text-based wizard-of-Oz study to understand the rules of practice in the space with respect to the devices (the array of lights), allowing us to integrate the rules of practice and the devices. We used an *in situ* speech-based training data collection study to integrate the system with the devices and rules of practice. We have done a preliminary evaluation with the live system and are currently conducting a longer evaluation comparing the speech interface to a graphical interface. At the end of each study in the design process, we use interviews and questionnaires to understand what kind of impact the system might have.

#### *Text-Based Wizard-of-Oz Study*

We began our design of Illuminac with a low-fidelity wizard-of-Oz study to better understand the lighting control domain including the kinds of lighting scenes used in the space and the types of commands used to refer to the lighting scenes.

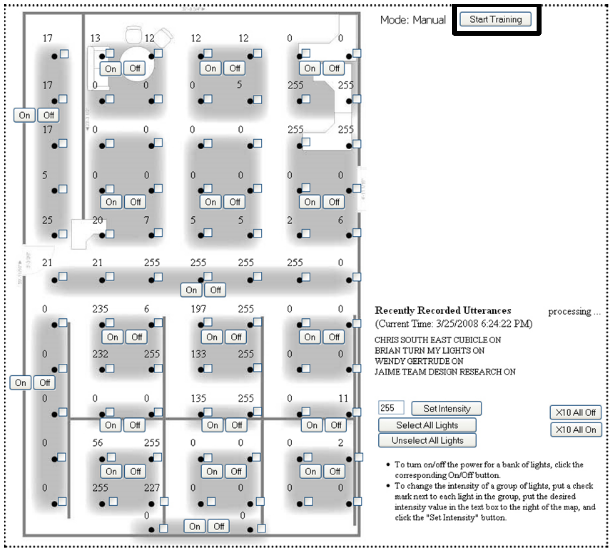


**Figure 2. Web-based graphical user interface used to manually configure the array of lights in the workspace.**

By low-fidelity, we mean natural language text input instead of speech input. Participants were asked to send the wizard (a researcher in the lab) an instant message whenever they wanted to change the lighting scene. The wizard would change the lighting scene using the web interface based on the participant's message. If the wizard was not available to change the lighting scene, the participants were asked to type the message they would have sent to the wizard into the web interface and then change the lighting scene themselves with the web interface. This way data could be collected even if the wizard was not at his desk. The GUI interface was similar to the one in Figure 2, but it also included a text box for entering the messages.

This study confirmed our hypothesis about the rules of practice that the lighting configurations are sometimes overlapping and are not all disjoint sets of lights. After the formal study concluded, some participants expressed the desire to continue being able to ask the wizard to change the lighting scene, but instead of sending instant messages, they wanted to ask the wizard with a spoken command. This provided anecdotal evidence that speech would be a good fit for lighting control in the space.

#### *Formative Training Data Collection*

After the low-fidelity wizard-of-Oz study, we collected two weeks of high fidelity training data to design and tune the learning algorithm, which estimates lighting scenes given a spoken command. The study included sixteen participants who were regular occupants of the space. Each participant was asked to record a command and demonstrate the desired system response as if they were training the system to understand their personalized commands. They were asked to complete the following three steps each time they wanted to change the lighting scene:

1. say their command to change the lighting scene;

2. type their name into the text box on the web page; and
3. change the lighting scene with the web interface.

We did not tell them when or how to change the lighting scene in the lab. The participants could use any of the eight microphones throughout the space to record their command. Then they used the web interface to demonstrate their desired change in the lighting scene. We followed the data collection with individual interviews and asked the participants to reflect on their lighting control preferences and their experience with the study.

We used this data to inform the design of the learning algorithm, which in turn enabled us to build and deploy a live system.

*Preliminary Evaluation of Deployed System*

We deployed Iluminac with ten of the twenty five regular occupants of the lab for one week. The system ran with live speech recognition on the audio from the microphones around the room, live training mode where the model was retrained after each training point was collected, and live running mode which applied a lighting scene when a user spoke a command into one of the microphones. We started with no training data and asked the participants to train the system on their commands again[1]. Participants were instructed to use the system whenever they wanted to change the lighting scene. The first time they used the system for a particular command they were asked to record a training data point. Subsequent times they were asked to test their commands, but recording more training points if the system did not respond as expected. When the participants tested a new command they recorded the accuracy of the results on a paper log next to the microphone. They recorded the accuracy of the system response by circling one of the following options:

| 4 | 3 | 2 | 1 | 0 |
|---|---|---|---|---|
| Correct | Partially Correct | Some Correct, Some Wrong | Nothing Happened | Wrong |

At the end of the study, the participants were asked to complete an anonymous web questionnaire about their experiences with the system. Figure 3 shows a sample of the commands collected in the study.

The average testing score plateaued between "correct" and "partially correct" when commands were tested with 1, 2, or 3 training points. Although the average score is closer to "partially correct" than "correct", when asked in the post questionnaire, "After the study is over, would you like to continue using the system?" eight out of ten participants responded *Yes*, and two responded *Maybe*—the options were *Yes*, *Maybe* and *No*. One of the participants who responded *Maybe* said the microphone was too far away, and he was too lazy to get to a microphone, which is a limitation in our experimental setup that could be overcome in a commercial deployment. Right now each cubicle with three people shares one microphone, which could be remedied by giving each user a microphone at their desk, making the microphone convenient to access. The other participant who responded *Maybe* tends to sit in the public area most of the time where the furniture moves around quite a bit, and he does not often use the same set of lights. In such an open space, a location-based speech approach would likely work better where users could say "lights on here." Such a location based approach could be implemented with distributed microphone array technology overhead, though such technology would not be desirable in the cubicle area for privacy reasons. With overhead microphones, users cannot control what is being recorded, but with desk microphones, users have the power to turn the microphone on their desk off.

[1]Many of the participants in this study also participated in previous studies and had already provided training points.

Brian: turn my lights on
Jack: my lights on
Chris: all on
Marcus: turn on my lights
Link: turn on my lights please
Jaime: team design research on
Joe: experiment lights on
Link: turn on all the lights
Chris: south east cubicle lights on
Link: turn on all of Chris's lights
Link: turn on the lights over the public space
Joe: presentation mode Chris: window lights off
Kevin: dim my desk lamp

**Figure 3. Sample lighting scene commands collected during the preliminary evaluation of the live Illuminac system.**

This study allowed us to evaluate how well the system worked and gave us preliminary qualitative data about the use of speech in the context of workspace lighting control, but did not allow us to accurately study the later. The act of recording the accuracy of the system each time a participant used the system interfered with the experience of using speech commands to control the lighting scene making it harder to study the participant's experience with the system. We made some small changes to the system based on feedback from this study and we are now running a longer study to understand the use of speech to control the lighting scene.

*Speech and Graphical Interface Comparison Study*

We are currently running a longer study with nine participants for an average of eight weeks. In this study we are looking at the experience of controlling the lighting scene with speech and a graphical user interface. The participants are asked to control the lighting scene with the graphical user interface for half the time, and with the speech interface for the other half of the study. We are not asking the participants to give *in situ* feedback on the system performance to avoid affecting the experience of using each of the interfaces. At the midpoint and at the end we will interview the participants about their experience with each interface.

## CONCLUSIONS

We have designed, implemented and deployed a natural language speech interface for workspace lighting control which we believe is an example in a larger problem space, namely natural language speech interfaces for controlling configurations of devices in the home or in workspaces. We have followed an iterative design process allowing us to integrate the devices, system and rules of practice. We used a lo-fidelity text based study to understand the configurations of lights and commands naturally used in the space, a formative training data collection study to collect high-fidelity data to inform the design of the learning algorithm. Once we deployed the live system we studied the accuracy of the system and are now studying the experience with the system as compared to a graphical user interface to the lighting control. The interviews at the end of each study have helped us understand how the system would be accepted, and the impact of the system on the users' behaviors and the behaviors of the entire lab.

## BIOGRAPHY

Ana Ramírez Chang is a PhD candidate at the University of California at Berkeley working with Professor John Canny.

## REFERENCES


1. Adura. Adura thchnologies, as of March 2008. http://www.aduratech.com.
2. DALI. Digital addressable lighting interface (dali), as of 24 June 2008. http://en.wikipedia.org/wiki/Digital-_Addressable_Lighting_Interface.
3. M. C. Mozer. Lessons from an adaptive house. In D. Cook and R. Das, editors, *Smart environments: Technologies, protocols, and applications*, pages 273–294. Wiley & Sons, 2005.
4. M. Weiser and J. S. Brown. The coming age of calm technolgy. In *Beyond calculation: the next fifty years*, pages 75–85. Copernicus, New York, NY, USA, 1997.
5. Y. Wen, G. J, and A. Agogino. Towards embedded wireless-networked intelligent daylighting systems for commercial buildings. In *Proceedings of the Institute of Electrical and Electronics Engineers Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*, 2006.