

# Out to Lunch: Further Adventures Monitoring Background Activity

Jonathan Cohen  
Human Interface Group  
Advanced Technology Group  
Apple Computer, Inc.  
present address:  
Interval Research Corporation  
1801 Page Mill Road  
Palo Alto, CA 94304  
Email: cohen@interval.com

## Abstract

The sound of keystrokes and mouse clicks generated by a number of computer users gives co-workers a sense of “group awareness”—a feeling that other people are nearby and an impression of how busy they are. “OutToLunch,” a software prototype, attempted to foster this sense of group awareness by using background sounds and an electronic signboard to inform physically dispersed or isolated group members about each other’s presence. Although the software was never formally tested, it was distributed within a group. The sound design and the messages displayed on the sign were changed in response to the group’s feedback, and two different methods for constructing background audio were utilized.

## 1 Introduction: Group Awareness

The liquid sound of keystrokes and mouse clicks generated by a number of computer users gives co-workers a sense of “group awareness”—a feeling that other people are nearby and an impression of how busy they are. When my group moved from a building with open cubicles to one with closed offices, we could no longer hear this ambient sound.

“OutToLunch,” a software prototype, attempted to restore this sense of group awareness by using background sounds and an electronic signboard (like a portable version of the electronic ticker-tape above Times Square) to inform group members about each other’s presence. Although the software was never formally tested, it was distributed within my group. The sound design and the messages displayed on the sign were changed in response to the group’s feedback, and two different techniques were used to construct the background sounds.

OutToLunch wasn’t intended to help users accomplish a specific foreground task. It ran in the background to try to foster a feeling of belonging to a group, even if group members were physically dispersed or isolated. It took advantage of the human ability to process information in the background by using sound. OutToLunch was designed to fit Buxton’s model of a “holistic” technology—one that can “mirror” human social, physical, and cognitive abilities [1].

## 2 Related Work

Gaver and Smith have shown that adding background sounds to a cooperative factory simulation improved users' understanding of the state of the activity [2].

Bly et al., have built a number of Media Spaces—virtual spaces using audio and video connections—to promote cooperative work between physically separated colleagues. Bly et al., were interested in facilitating both formal and informal collaboration [3]. Gaver, et al., employed background sounds to warn users that they were being viewed remotely by other users over another media space, an interoffice video network called RAVE. Gaver, et al., meant for these sounds to mirror—not mimic—the real world sights and sounds generated when people walked past other people's offices and glanced in Gaver et al., [4]. One of the most important concerns in the design of RAVE was the issue of privacy. User's views and reactions were collected and accommodated in the design [5].

Cohen employed background sounds to inform users about file sharing, an autonomous background activity. Users tested a number of different mappings of background sounds to file sharing events. Users found the sounds informative, but often described them as obtrusive or annoying [6, 7]. In an alternative graphical display, icons moving slowly up one side of the screen represented file sharing activity. Most users found this too distracting, but one found it useful on a second monitor located in his visual periphery. There, the motion was just enough to alert him, but not enough to distract him [7].

Jackson and Francioni mapped computational events in a parallel program to musical notes played in different timbres. Programmers were better able to perceive more information about program behavior using graphics with sound than with graphics alone. Notably, the resulting music, made up from the sounds of many small events, was pleasant to listen to Jackson and Francioni [8].

## 3 The Data

OutToLunch counted each user's keystrokes and mouse clicks and kept track of the pixel distance each user's mouse was moved. This data went to a server that calculated the total keystrokes, mouse clicks, and mouse distance for the group. The total, updated every thirty seconds, was displayed to users in either text or audio form. The text appeared on an electronic signboard, mounted in a public area that group members often walked through. A background application, running on a computer in each group member's office, played the audio.<sup>1</sup>

OutToLunch displayed group totals, not data about individuals, since this seemed enough to give a sense of group awareness. This also ensured that the details of individuals' work habits were kept private (especially given the public location of the sign).

## 4 The Sounds

The auditory display ran on Macintoshes with no special audio hardware. To simulate the real ambient sound, OutToLunch played a prerecorded keystroke or mouse click sound for each keystroke

---

<sup>1</sup>Here is the story behind the name "OutToLunch." During the day, a user would hear the sounds informing him that other members of the group were present. Around noon-time, if the sounds ceased, the user might guess that everyone else had gone out to lunch and left him behind.

or mouse click input by someone in the group. At playback time, OutToLunch chose randomly from different prerecorded sounds of the keystroke or mouse click type. Only the total count of keystrokes and mouse clicks over a thirty-second interval was recorded, not the exact timing of each input keystroke or mouse click, so OutToLunch played the prerecorded sounds back at random offsets within the interval—preserving the temporal “density” of the input, but operating on a time lag.

This mapping of small events, keystrokes and mouse clicks, to sounds is close to the technique used by Jackson and Francioni. There is a small difference in using sampled sounds instead of musical notes, and a deeper difference because the Jackson and Francioni mappings are meant to be heard in the foreground, while OutToLunch is meant to be heard in the background.

The mouse distance was mapped to audio by adding one prerecorded “mouse-rolling” sound to the mix for every few hundred pixels traveled during a thirty-second interval.<sup>2</sup> As above, there were several prerecorded sounds to choose from, and they played back at random offsets within the interval. Users noticed the mix was less obtrusive when there were more mouse-rolling sounds, so I reduced the mouse distance mapped to a single mouse-rolling sound to 100 pixels.

I believe the mouse rolling sounds made the mix more pleasant because they had a white-noise component, and a longer duration than the clicking sounds.<sup>3</sup> The rolling sounds gave the mix a more even quality which the clicking sounds could rise out of; the clicking sounds stood out too sharply without a white noise mask.

Nevertheless, everyone found OutToLunch annoying to listen to for very long (except me). I suspect this was not only because of aesthetic dissatisfaction, but also because there wasn’t much information in the sounds. Compared to the real keystroke-and-mouse click ambiance, there were three flaws in the OutToLunch sounds: they were louder but had lower resolution; all directional information was lost; and users couldn’t tell whether the sounds echoed their own activity or someone else’s (unless a user wasn’t currently typing, and therefore the sounds had to be echoing someone else’s activity).<sup>4</sup>

## 5 The Sign

The sign was mounted in a highly-frequented public space in an effort to make it less obtrusive for individual users. They might notice it as they walked down the corridor, but it wouldn’t disturb them in their offices. This location was chosen as a kind of group analog to a monitor placed in a single user’s visual periphery (as described above).

The sign displayed messages of the form “In the last 15 minutes, 3 people on the *X* team generated 437 keystrokes, 134 mouse clicks, and .001 mouse miles (at 72 dpi).” Although users found the messages funny at first, they grew tired of them. However, two users told me they still looked at the sign first thing in the morning, before they started typing, to tell whether other people in the group had already arrived.

Since OutToLunch provided so little information, I asked people in our group what else it could reveal without making them uncomfortable. People did not want the name of the application they

---

<sup>2</sup>Users weren’t able to intuitively recognize the mouse rolling sounds. But they were able to remember the mapping of the sounds to mouse distance once the sounds were explained.

<sup>3</sup>Mouse rolling sounds lasted from 0.5 to 0.66 seconds, keystrokes and mouse clicks from 0.2 to 0.3 seconds.

<sup>4</sup>This flaw could have been fixed—as OutToLunch converted the activity totals to sound on a particular user’s machine, it could have subtracted the activity generated by that user.

were currently using displayed, nor did they care for messages like “*X* is not currently using his [her] computer.” However, no one objected to messages saying “*X* is active,” or more accurately, “*X*’s machine is active.” Did “*X*’s machine is active” also imply that *X* is busy, and not open to interruption? Everyone in the group said they were actually most interruptible when they were typing at the computer. But they also said they were least likely to interrupt someone else if they saw that person was busy typing!

## 6 Out to Lunch Revised

The messages on the sign were changed in response to this feedback, but the sound design had to wait until I had the good luck to collaborate with Michael Brook, who is a musician, composer, and record producer. In this version of OutToLunch, each person in the group was represented by a specific theme (Blattner has suggested using musical themes to encode information with sound [9]). A steady guitar sound (a “drone”) played continuously whenever anyone in the group was typing. Once every thirty seconds, for a period as long as a particular user continued typing, the user’s theme was mixed in with the drone. The drone and themes were composed and realized by Michael Brook, who also specified how they should be mixed together.

OutToLunch used several methods to keep its mix unobtrusive. The drone—an arhythmic, low-pitched, low-volume, seamless loop of solo guitar music lasting 13 seconds—established an even ambiance. When the first person in the group started typing, the drone faded up from silence over a nine-second interval.<sup>5</sup> OutToLunch mixed in themes only when the drone reached its full volume. If, over time, all group members ceased typing, the drone faded back to silence over a nine-second interval. The themes lasted from five to ten seconds, and all had a fading, echoing quality, and a soft attack. To keep the mix from sounding busy, a given user’s theme could be played only once per thirty-second interval. To add variety to the mix, each time a theme played, it started up at a random offset within the interval; OutToLunch seldom repeated itself.

## Conclusions

OutToLunch was not meant to be a software tool for accomplishing a particular task. Its purpose was simply to promote good feelings between colleagues, and that seems to me to be a worthwhile goal regardless of the success or failure of OutToLunch. The program required only a low-bandwidth network, connecting computers capable of playing back eight-bit sound. Nonetheless, it brought the group to confront issues of privacy and what exactly it meant to be “busy.”

Users have commented on how pleasant-sounding the new OutToLunch is. Since there were only six people in the group, no one had trouble associating a theme with the person it represented (users each chose their own theme from a set provided by Michael Brook). Unlike the earlier version, users get feedback about their own keystroke and mouse activity, and perhaps a sense of when other users might be open to interruption. But having just released the new version, I don’t know whether users will find it pleasant to listen to after many hours, or whether it will succeed in fostering a sense of group awareness.

The following anecdote is mildly supportive, however, and Bly tells a similar story [3]. I was reading an article one morning and a came across a minor point that was unclear. Just at that

---

<sup>5</sup>A longer fade would have been less obtrusive, but would have increased time lag.

moment I noticed that Tom's theme had been playing, so I decided to call him on the phone and discuss the point. Hearing the theme helped crystallize the idea of calling Tom. But it also meant it was very likely Tom would pick up his phone, so I wouldn't have to deal with a tedious voice mail system. This interaction seemed as lightweight as if T had been in the office next door, and I had popped in for a quick chat, though in reality, T's office was on a different floor on another side of the building.

One conclusion one should be careful **not** to draw is a general one that motifs work better than everyday sounds for this kind of application. The everyday sounds in OutToLunch had two strikes against them: they were less informative than the motifs, and they were designed by computer scientist rather than a sound designer. A good sound designer could redesign the simulated keystroke sounds to be pleasant to listen to and allow users to tell which keystrokes belonged to which person.

Here are some conclusions that can be drawn from experiences with OutToLunch.

- A steady sound can provide a base which sharper sounds can rise out of, without making the overall ambiance much more obtrusive. To an extent, this was even true in the first version of OutToLunch.
- Sound designers, musicians, and composers know many useful methods for creating a pleasant ambient mix. Just as it's now normal to work with graphics designers to develop GUIs, it should be normal to work with sound designers to develop audio UI's.
- Privacy issues come up even in a low bandwidth system like OutToLunch. As was true in the RAVE work, there is a line between privacy and group awareness that might best be drawn by the people affected by the technology [5].
- People may be open to interruption at unexpected times: some users valued typing less than talking.

## 7 Future Directions

The method of constructing a background ambiance by mixing together bits of sounds representing many small events may be a useful one. This kind of constructed sound could be mixed together with an overall flow of sound representing state to generate a pleasant, informative ambiance. For example, a much less literal simulation version of OutToLunch could be created, where each keystroke was mapped to some short musical sound (after Jackson and Francioni) with an underlying drone representing whether people were working at home or at the office.

Following Gaver [10] or Blattner [9] more information could be encoded into the ambiance. For example, the pitch of the drone could be raised or lowered depending on the overall group activity level.

OutToLunch could combine input from an "office-presence" sensor with keyboard and mouse activity, to let group members know that others were present, but not typing—and therefore probably not open to interruption.

OutToLunch doesn't scale up. The ambiance would not remain unobtrusive if the group got much bigger than six people, and with a larger group size, users would have trouble associating a theme with a particular group member. Would users ever want to have a sense of group awareness

with more than a few other people at a given time? What are some mechanisms that might make OutToLunch work for a larger group? One possibility, for example, is that all the themes for a particular subgroup of people could use a similar timbre. A user could increase the volume of themes within the mix for people or subgroups the user particularly wanted to know about.

## Acknowledgments

Thanks to Tom Erickson for the conversation that originated the idea and name of OutToLunch; to Tom Erickson, Gitta Salomon, Joost Kemink, Anne McClard, and Sally Grisedale for their patience working with different versions of the program; and to Michael Brook for useful discussions and the composition, sine qua non. Thanks to Apple Computer, Inc. for permission to publish.

## References

- [1] Buxton, W. "The Three Mirrors of Interaction: A Holistic Approach to User Interfaces." In *Interacting with Virtual Environments*. New York: Wiley, 1994.
- [2] Gaver, W., and E. B. Smith. "Auditory Icons in Large-Scale Collaborative Environments." In *Interact '90: Proceedings of the IFIP TC 13 Third International Conference (1990)*: 735-740.
- [3] Bly, S. A., S. R. Harrison, and S. Irwin. "Media Spaces: Video, Audio, and Computing." *Communications of the ACM* 1 (1993): 28-46.
- [4] Gaver, W., T. Moran, A. MacLean, K. Carter, and W. Buxton. "Realizing a Video Environment: EuroPARC's RAVE System." In *Proceedings of CHI '92 (1992)*: 27-37.
- [5] Bellotti, V., and A. Sellen. "Design for Privacy in Ubiquitous Computing Environments." In *Proceedings of the Third European Conference on CSCW*. Dordrecht, Netherlands: Kluwer, 1993.
- [6] Cohen, J. "Kirk Here: Using Genre Sounds to Monitor Background Activity." In *Adjunct Proceedings of InterCHI '93 (1993)*: 63-64.
- [7] Cohen, J. "Monitoring Background Activities." In *Auditory Display: Sonification, Audification, and Auditory Interfaces*, edited by G. Kramer, 253-266. Saanta Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading MA: Addison-Wesley, 1994.
- [8] Jackson, J. A., and J. M. Francioni. "Synchronization of Visual and Aural Parallel Program Performance Data." In *Auditory Display: Sonification, Audification, and Auditory Interfaces*, edited by G. Kramer, 291-306. Saanta Fe Institute Studies in the Sciences of Complexity, Proc. Vol. XVIII. Reading MA: Addison-Wesley, 1994.
- [9] Blattner, M., D. Sumikawa, and R. Greenberg. "Earcons and Icons." *Human-Computer Interaction* 4 (1989): 11-44.
- [10] Gaver, W. "Synthesizing Auditory Icons." In *Proceedings of InterCHI '93 (1993)*: 228-235.