
Auditory Priming for Upcoming Events

Timothy Sohn

Nokia Research Center
955 Page Mill Road
Palo Alto, CA 94304 USA
tim.sohn@nokia.com

Leila Takayama

Willow Garage
68 Willow Road
Menlo Park, CA 94025 USA
takayama@willowgarage.com

Dean Eckles

Nokia Research Center
955 Page Mill Road
Palo Alto, CA 94304 USA
dean.eckles@nokia.com

Rafael Ballagas

Nokia Research Center
955 Page Mill Road
Palo Alto, CA 94304 USA
tico.ballagas@nokia.com

Copyright is held by the author/owner(s).
CHI 2009, April 4 – 9, 2009, Boston, MA, USA
ACM 978-1-60558-246-7/09/04.

Abstract

Psychologically preparing for upcoming events can be a difficult task, particularly when switching social contexts, e.g., from office work to a family event. To help with such transitions, the audio priming system uses pre-recorded audio messages to psychologically prepare a person for an upcoming event. In this system, audio priming is being used to prepare a person's state of mind to improve one's sociability in the upcoming social context.

Keywords

Auditory user interfaces, auditory priming, event priming, preparation, context-aware, mobile

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

The many events and meetings in our daily lives can be overwhelming. The process of preparing for an event may involve gathering information about the event, knowing who will be attending, and being in the appropriate mindset for the event. Sufficient preparation is often time consuming and requires mentally transitioning from one event to another. Transitions such as moving from one project meeting to the next, from a professional mindset to a personal

one, or from one social group to another, present difficulties for many individuals. For example, a busy executive may have many consecutive meetings with little time in between to prepare. A working mother may have difficulty leaving her work mindset behind as she transitions from several work meetings to her child's soccer game. New tools could address these difficulties and support preparing for upcoming events.

Related Work

Information gathering for upcoming events is not a new idea [2], but many interface design questions remain open. The audio priming system differs from prior work in (1) its use of an audio interface optimized for the demands of mobile, inter-event contexts and (2) its inclusion of information and modes of presentation designed for more than explicit information transfer.



Figure 1. Example use case scenario for the audio priming system

Personal information management (PIM) systems are often used to support preparation for upcoming events by tracking data needed for an event, such as the attendees, location, and additional notes for an event. Microsoft Outlook offers these capabilities as well as group sharing that give a person access to other types of personal data from other people using the same system. PIM systems offer immediate access to information used in preparing for an event. However, these

existing systems are very visually intensive and often difficult to access during crucial transition periods.

An optimal time to prepare for an event is immediately before the event is to take place. For many people this is a period of multiple demanding transitions: There is the physical transition where a person is mobile while traveling from one event to another, and a cognitive transition from the current mindset to the appropriate one for the next event.

When these transitions happen in mobile scenarios, visual interfaces are often not optimal. For example, attempting to read information from a mobile phone while driving is both dangerous and illegal in some places. This poses a problem for people who are trying to prepare while in transit. One possible solution to this problem is to deliver the information through the auditory perception channel instead of the visual one. This could be through a mobile phone's speaker, headset, or car stereo system. Unlike text and images that require reading and visual attention, an audio presentation can be more passively used and more easily ignored (e.g., when the driving situation required very focused attention).

While audio is likely the most appropriate medium for a driving scenario, the system needs to know what information to offer through the audio channel (Figure 1). A naïve approach would be to read the names of the attendees, location of the event, and any additional notes to the user. This would be helpful for the user to hear the content of the information, but may not be the best approach for preparing that person for the upcoming event.

Rather than presenting information to explicitly learn or remember, we take a more holistic approach by presenting data through pre-recorded audio messages in the voices of the likely attendees to help prime the user for the event. A simple PIM system could portray relevant information, such as all the attendees for an event, through an audio channel. Our system leverages the natural voice of the attendees to make similar data more engaging to activate memories and allow the user to anticipate the interaction.

Also unlike current technological solutions, the audio priming service can present additional content in the voices of the people the user is likely to encounter soon, thereby psychologically priming the user to be in the cognitive and emotional state for the particular social encounters ahead. This can be content directly related to the expected focus of the event, but it can also be less related content that is likely to, e.g., increase enthusiasm for the event and seeing the likely attendees.

Priming

In psychology, “priming” refers to making certain categories and ways of thinking temporarily more accessible for the near future [5]. In many cases, priming can occur below the level of conscious experience; see [1] for a review. By leveraging this capacity of the human psyche to be influenced by preceding stimuli, the audio priming system aims to positively influence upcoming events by activating appropriate representations and mindsets for those upcoming events. Our goal is to allow a person to hear all the necessary information while helping him transition both cognitively and emotionally for the upcoming event.

By using familiar voices to prime users, this system leverages natural human abilities to identify individuals by their voices and to be more psychologically prepared for encounters with them. As shown in brain imaging studies, recognizing voices of familiar individuals activates the fusiform gyrus of the brain, which is part of the visual system that is used in face recognition [8]. This suggests that being primed with familiar voices of particular individuals will prepare the user to more readily recognize and interact with those people at the upcoming event. Furthermore, hearing a familiar voice activates brain activity in areas implicated in episodic memory and emotional salience [6], which suggests that memories and emotional responses to those individuals will be primed in the user’s brain before arriving at the upcoming event.

System Design

The system takes input from various data sources with calendar information. This data is analyzed to extract the time of the event and a list of attendees. Based on this list, the system searches a person’s personal information database to find the phone number for each of the attendees. Once the phone numbers have been collected, each number is dialed to get that person’s voice mail. The voice mail messages are saved as sound files either on the device or a server. When the upcoming calendar event is within a short period of time (e.g., 15 min), the system plays back the voice mail prompts for each of the attendees at the upcoming event. Figure 2 presents a system diagram of this entire process. Each of its components are discussed in the following sections.

Calendar Data Module

The calendar data module takes a collection of calendar events from multiple data sources (e.g., Facebook, Evite, Google Calendar, Microsoft Outlook). For each of these data sources, the system extracts upcoming events along with a list of attendees for each event. The attendee list can be in the form of names, email addresses, or phone numbers. The identity of each attendee is passed on to the audio priming system to prepare for delivery.

Gathering Voice Content

The voice content module is responsible for obtaining rich personal voice messages for each of the attendees at an event. Once the audio priming system has resolved the identity and phone numbers for a list of attendees, the voice content module attempts to gather rich voice messages from the following sources: (1) recent voicemail left by the attendee, (2) voicemail

greeting of the attendee, (3) recorded initial 'greeting' portion of telephone conversations, (4) voice conversion technologies to make the computer voice of automatically generated text sound like the attendees' voices. Each of these sources adds a personal dimension to an attendee's name. Typical voicemail greetings and messages start with "Hi, this is Joe..." Recording these initial seconds of a voicemail message or greeting is an effective way to prime the user about an upcoming event in each attendee's natural voice.

Providing Audio Priming Service

Based upon contextual data (e.g., time of day, GPS-tracked location, proximity to car stereo), the priming service may be proactively offered while en route to the upcoming event. For example, if the system detects that the user is driving and an event is within 15 minutes, it could begin the audio priming process. Alternatively, the user may manually activate the priming service.

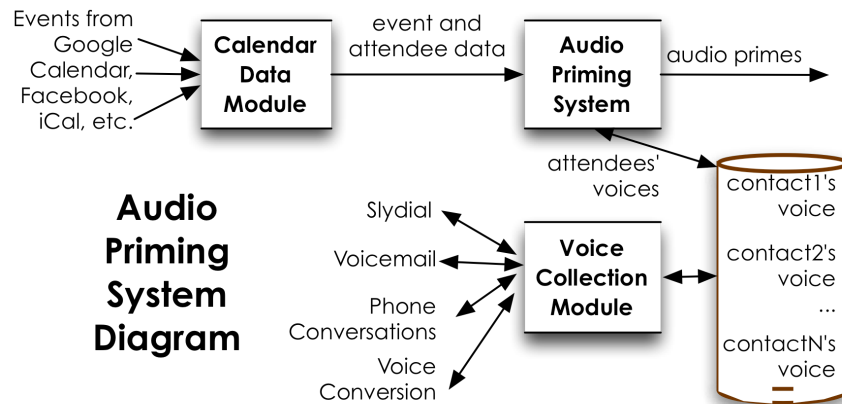


Figure 2. Data flow system diagram from the audio priming system

Regardless of how the interaction is initiated, the audio priming service provides the audio signal to be played through the currently available audio output hardware in the user's vicinity, e.g., mobile device, headphones, car speakers, laptop speakers, etc. A benefit of audio is that it is an ambient form of information delivery that will less directly interfere with ongoing important tasks such as traveling to the upcoming event.

Audio primes provided by this service may be as simple as a list attendees' names provided by the voice collection module or may be much more complex. Some examples of more complex audio primes include: conversational topics that the attendees are likely to raise, reminders about the significant others and family

members in the attendees' lives, news items that relate to the attendees' interests and activities, or blog headlines and status messages from attendees' recent online posts.

Current Implementation

The current work is aimed at supporting those users who have very large social networks, who are forgetful of names, who are shy in social situations, or who would otherwise benefit from the warm-up for upcoming social events.

We have implemented a proof-of-concept auditory priming system, allowing users to hear the names of upcoming event attendees in their own voices. In its current form, the system uses shared calendar information from Google Calendar and Facebook Events to identify upcoming events and attendees. Both of these event systems incorporate information about who will attend the event, which helps the system know which attendee voice messages to gather.

Neither of the event systems directly exposes a phonebook interface. As a result, our system performs entity resolution with a person's name or email address and the user's contact database. The contact database can be the user's personal phonebook on their device or other cloud-based services (e.g., Ovi Contacts or Google Contacts). If an appropriate phone number cannot be found, the attendee's voicemail is not included in the priming process.

The current system is implemented on the Nokia N95 mobile phone. Once the system recognizes an upcoming event and resolves phone numbers for the attendee list, each phone number is dialed to record

the attendee's voicemail greeting. We currently use Slydial [7] to access each attendee's voicemail message, record the greeting, and compile the voice recordings into an audio prime sequence for the user.¹ The audio gathering process only needs to be done once for each attendee because the system saves the associated audio prime for later use. When the event is within 15 minutes, the phone prompts the user to playback the audio prime sequence. The audio prime sequence can be played on a Bluetooth-enabled audio device (e.g., headset or car stereo).

Although our voicemail gathering process is automated, clipping the audio recordings to create the audio prime sequences is still a manual process. In the future, we would like to explore methods to automate this process by detecting the voicemail prompts to appropriately trim our recording or by using other types of personal voice messages as described earlier. We also plan to incorporate other event management sources such as Evites and Microsoft Outlook.

Future Work

To empirically evaluate our hypothesis that audio primes will improve users' readiness for upcoming social events, we have designed an experiment to be run in an ongoing field deployment of mobile information services. More specifically, this study will examine the cognitive and affective consequences of using different versions of the audio priming system. These variations will be generated from such independent variables as including audio prime content (e.g., saying own name vs. user's name), audio prime

¹ Slydial lets the system call straight to a person's voicemail without ringing the other person's phone.

source (e.g., using own voice vs. user's voice vs. text-to-speech), or priming modality (speech vs. text).

Our next step will be to pilot the system among office workers, who should by hypothesis benefit from audio primes for their upcoming meetings throughout the workday and while transitioning from their work life to personal life. Based upon feedback on this initial system, we will then decide whether and how to incorporate more complex audio primes for a larger field deployment.

Implications for HCI

The auditory priming system provides a starting point for delving into interaction designs that support psychological preparedness and social context switching. This is contrasted with existing systems and research projects that tend to support ongoing task performance and cognitive task switching [3][4]. Our ongoing investigations in this space are aimed at widening the view of human-computer interaction research, informing the design of systems that take advantage of the psychological priming mechanisms – in this case for improving future face-to-face interactions.

Recommendations

Our goal in sharing this work-in-progress is to stimulate related research activities around this design space of anticipating upcoming social events and how user interfaces might better support users' psychological preparedness for those events. The current work focuses upon the specific situation of supporting

individual users, who are going to be attending social events in the near future. This is merely one instantiation within a much broader design space of priming systems. Further investigations in this space could also address broader issues of how priming could be used to positively influence people's everyday lives.

Citations

- [1] Bargh, J. A. What have we been priming all these years? *European Journal of Social Psychology* 36, (2006), 147-168.
- [2] Burkey, C. and Ho, D. Advanced information gathering for targeted activities. Patent number 6845370 (2005).
- [3] Card, S. K. and Henderson, A. A multiple, virtual-workspace interface to support user task switching. In *Proc. CHI 1987*, ACM Press 1987), 53-59.
- [4] Czerwinski, M., Horvitz, E. and Wilhite, S. A diary study of task switching and interruptions. In *Proc. CHI 2004*, ACM Press (2004), 175-182.
- [5] Manstead, Antony S. R. and Miles Hewstone (Eds). *The Blackwell Encyclopedia of Social Psychology*. Blackwell Publishing, 1996.
- [6] Shah, N. J., Marshall, J. C., Zarifis, O., Schwab, A., Zilles, K., Markowitsch, H. J., and Fink, G. R. The neural correlates of person familiarity: A functional magnetic resonance imaging study with clinical implications, *Brain* 125 (4), (2001), 804-815.
- [7] Slydial. <http://www.slydial.com>
- [8] von Kriegstein, K., Kleinschmidt, A., and Giraud, A. Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia, *Cerebral Cortex* 16 (9), (2006), 1314-1322.