Quick access to information can save lives. This is especially true for first responders like the police, fire department, or emergency medical technicians. First responders commonly receive information to aid them in their field duties from a radio dispatcher, if one is available, over a two-way radio. Recently, in-vehicle laptops have become an option for direct data access over wireless networks. But the laptop systems have some limitations. For example, they don’t offer hands-free or eyes-free interfaces. Furthermore, they often use advanced wireless technologies for data transmission, such as Cellular Digital Packet Data systems. Such technologies aren’t always available, especially not in poor rural areas.

In Alabama, where we developed the Voice Law Enforcement Tactical System (VoiceLETS), some financially constrained rural counties can’t offer police and other first responders access to radio dispatchers. Even in counties that do, the dispatchers are often overworked, which can limit access to them. The same budget constraints that restrict the number of dispatchers and increase their workloads also keep many agencies from implementing in-vehicle laptop systems.

Yet lack of access to information can put first responders at a potentially fatal risk. For example, figure 1 shows an officer approaching a stopped vehicle from the rear. He will be safer if he knows the car’s status before he addresses the driver. If the vehicle is stolen, he can prepare for an arrest rather than a routine speeding ticket.

Auburn University and the University of Alabama developed VoiceLETS in a collaborative research effort sponsored by the Southwest Alabama Integrated Criminal Justice System to address this issue for rural law enforcement agencies. We focused on fielded and established technologies, using the existing cellular infrastructure to transport data over voice channels. VoiceLETS gives police real-time, voice-based access to driver and vehicle information in a low-cost system design.

**EDITOR’S INTRODUCTION**

As most Pervasive Computing readers know, highly portable devices such as high-end cell phones and networked PDAs can offer first responders like the police timely access to mission-critical information. However, lack of integration with enterprise databases and limited input capabilities—thumb keyboards, for example—have hampered their deployment, especially in hands-busy and eyes-busy environments. Still, systems like Vocera’s communication badges for in-hospital voice dialing, which we covered in the July 2003 issue, show that voice integration of back-end recognition servers, enterprise databases, and wireless infrastructure can overcome obstacles to successful deployment.

In this issue, we cover the Voice Law Enforcement Tactical System developed by Juan Gilbert and his colleagues in conjunction with the Auburn Police Department in Alabama. VoiceLETS makes traffic stops safer for patrol officers through a cell phone voice interface to driver and vehicle data. A back-end speech recognition server forms database queries, and the system returns search results to the officer without need of a human dispatcher. Then officers know if they are dealing with an ordinary traffic stop or a potentially dangerous confrontation over a stolen car. Because many rural counties cannot afford the necessary dispatchers, automated pervasive systems are their only viable alternative.

VoiceLETS has an architecture that is broadly similar to the Vocera voice-dialing system but is geographically distributed using cell phones rather than 802.11-based communication badges. I believe we will see this basic pervasive system architecture replicated in many future applications. The evolving VoiceXML standard will support its deployment as well.

—Vince Stanford

**BACK-END ENTERPRISE DATA**

VoiceLETS accesses data in the existing Law Enforcement Tactical System (http://care.cs.ua.edu/lets.aspx). LETS is a secure Web-based search engine designed to provide personal and vehicle information to law enforcement and criminal justice agencies. It integrates access to several state databases and currently includes photographs, addresses,
personal characteristics, and driver and criminal histories from the state's motor vehicles and corrections departments. The University of Alabama’s CARE Research and Development Laboratory (http://care.cs.ua.edu) developed LETS with the goal of making it available to all qualifying agencies over the Internet, including patrol officers using mobile data terminals.

VoiceLETS gives officers voice access to LETS data over a cell phone. Officers can search using a vehicle tag number or a personal driver's license or social security number. For vehicle queries, VoiceLETS will return the owner’s name, the vehicle’s make, model, year, and color. For queries on individuals, the system will return the driver’s name, address, weight, height, gender, race, number of warrant records, and the number of protection orders. In either case, officers can get critical information in less than a minute without a human dispatcher.

**CALL-PROCESSING ARCHITECTURE**

Figure 2 shows the basic call-processing infrastructure. A cell phone call is transmitted to a Session Initiation Protocol server. (SIP is a signaling protocol standard for voice over IP.) The SIP server relays the data to the voice server, which uses a voice interpreter and controller to handle automatic speech recognition and speech synthesis.

The voice server processes VoiceLETS scripts that interact with the officer and connect to the LETS database, housed in the Alabama Administrative Office of Courts.

**APPLICATION DESIGN**

VoiceLETS is based on the Voice Extensible Markup Language, a dialogue design language that supports development of voice-enabled Web pages with optional visual displays. VoiceXML is an open standard so a variety of platforms can interpret Web pages that use it. It has become the language of choice for many voice applications.

Key VoiceXML concepts include sessions, applications, dialogs, forms, menus, fields, grammars, subdialogs, variables, events, and scripting. These functions give developers complete control over application interactions through spoken dialogs. The application prompts the end user to respond in the context of the prompts. VoiceXML includes a voice browser that accepts audio input (voice or keypad dial tones) and audio output (computer-synthesized or recorded).

VoiceXML shields developers from the low-level implementation details of speech recognition, audio management, and text-to-speech output. However, the design of these systems can also be extremely challenging given the temporal nature of speech and the difficulty of representing hierarchical menu structures.

**Voice interface design issues**

Designing an effective voice interface for law enforcement officers in the field required a new application architecture and view of usability.

Graphical interfaces let users look or
PERSPECTIVE

Application systems often use problem, task-oriented speech-recognition systems suffer from machine speech-recognition errors. To mitigate this problem, task-oriented speech-recognition systems often use grammars that narrow acoustic searches to specific domains. These grammars specify words and phrase structures that the system will recognize. This reduces the chance of errors for in-domain speech, but it requires users to speak within the task grammar. If a user says something outside the active grammars, the system won’t recognize it correctly. These misrecognitions are called out-of-grammar or out-of-vocabulary errors.

Further, systems can suffer from recognition errors due to background noise, accented speech, and compression applied by telephony codecs. For example, if the system is in listening mode and a loud noise occurs in the background, it might force-fit the background noise to the in-grammar speech acoustics, particularly in systems without voice activity detectors to reject nonspeech noise.

VoiceLETS addresses these problems through a user-centered design consisting of three major VoiceXML dialogs: authentication, query, and results. Figure 3 shows the relationships among the VoiceLETS dialogs. The boxes represent dialogue between the user and the system. The diamond represents a decision dialog—in this case, authentication of previously accepted input. The lines and arrows represent the conversation flow. We implemented each VoiceLETS dialog to include basic components, such as small dialogs or functions that we could reference in multiple places. Basically, this is an object-oriented approach to voice application development where the objects are dialogs or grammars.

Constructing query grammars

When an officer dials into VoiceLETS, the system first asks for a five- or six-digit access officer authentication code. Each officer has a unique code.

After authentication, VoiceLETS opens the query dialog. We designed the query dialog to let law enforcement officers speak naturally when requesting driver and vehicle information. Officers can use several common preambles to the actual number they are referencing in a transaction—for example, “Search for driver’s license ID 1 2 3 4 3 1 0,” “Find social security number 7 1 6 0 9 3 8 7 1,” or “Look up tag 3 4 1 union George.” This gives VoiceLETS a degree of flexibility in handling the variety of query forms that officers commonly use. It also reduces training time by relieving officers of having to remember a rigid command syntax.

We incorporated the police spoken alphabet into the query grammar to model the way officers normally speak to dispatchers. This approach also increases the voice-recognition rate by eliminating near homophones in the grammar and using more phonemes per letter. Letters such as B, D, and P are easily confused, particularly over compressed telephony channels.

The two-way radio design metaphor was also useful in reducing errors from out-of-grammar input and background noise. VoiceLETS turns off VoiceXML’s barge-in function—a standard telephony feature that lets users interrupt the system, effectively supporting speaking and listening simultaneously. This makes the system more vulnerable to noise-induced errors—a serious problem for VoiceLETS because law enforcement officers aren’t insulated from background noise on the highway. VoiceLETS restricts officers from speaking until they hear the familiar beep sound of a two-way radio.

After the officer speaks the query, VoiceLETS confirms the request. Once confirmed, VoiceLETS submits the query to the LETS database and enters the results mode.

Results dialog

VoiceLETS automatically distinguishes between driver and vehicle queries, returning different data for them. Officers can replay spoken information by saying “repeat.” They can request the spelling of the name associated with a driver’s license ID number, social security number, or vehicle tag by saying “spell.” While searching for a vehicle’s tag information, an officer can say “owner” and the system will perform a license ID search on the vehicle’s owner.

Officers can return to the query mode by speaking “start over” at any time. Officer safety makes it necessary to alert officers early when a driver or

Figure 3. VoiceLETS interface dialog flow diagram showing authentication, query, and response system components.
vehicle has a warrant or protection order. VoiceLETS accomplishes this by playing a siren sound before speaking any query results associated with a warrant or protection order. If VoiceLETS begins speaking the name and associated information without playing the siren, the officer immediately knows that no warrants or protection orders are outstanding.

**USABILITY EXPERIMENTS AND ANALYSES**

We conducted an experiment to evaluate VoiceLETS’ effectiveness and usability. The goals were to determine the voice interface’s effectiveness for actual law enforcement/police officers and to identify possible improvements in its design and implementation. In addition, we queried officers’ satisfaction with the system as a measure of its usability.

**Methodology**

We selected 26 officers from the Auburn Police Department to participate. None had prior exposure to VoiceLETS. We gave each participant an instruction sheet outlining the query tasks and the steps required to complete them. The instructions contained a brief explanation of the system, specific goals of the experiment, and the task scenario.

Each participant completed the experiment using VoiceLETS to enter specified queries and retrieve the results. The experiment consisted, first, of authenticating access codes and, second, of issuing queries and retrieving results for

- a driver’s license number without warrants, warrants, or protection orders;
- a vehicle tag number;
- a social security number; and
- a driver’s license number with a protection order outstanding.

The last task is separate because it includes a distinguishing siren prompt to warn of outstanding protection orders or warrants.

Each participant worked alone using a telephone. Some participants completed the experiments inside while others moved outside the building to a parking lot where the external noise and disturbances more closely resembled real-life scenarios. We recorded each experiment electronically. This let us measure the time taken to complete the experiment, total number of tasks successfully completed, and number of attempts taken to complete each task.

After the experiment, we asked participants to complete a questionnaire. It included three biographical questions and six questions to evaluate their experience:

- To what extent did the system understand what you said?
- To what extent was the system’s speech easy to understand?
- To what extent did you know what to say in response to the system’s prompts?
- To what extent did the system behave the way you expected?
- Did you complete the task and get the information you needed?
- After using this system, how likely are you to use it or a similar system again?

Participants answered questions by rating them on a nine-point Likert scale, where 1 indicated never, unlikely, or difficult and 9 indicated always, likely, or easy.

The questionnaire also welcomed officer comments to evaluate their experience:

**Average user satisfaction with VoiceLETS correlated positively with the success rate (Pearson’s r = 0.478). The overall success rate was 92.31 percent.**

- To what extent did the system understand what you said?
- To what extent was the system’s speech easy to understand?
- To what extent did you know what to say in response to the system’s prompts?
- To what extent did the system behave the way you expected?
- Did you complete the task and get the information you needed?
- After using this system, how likely are you to use it or a similar system again?

**Results**

We used two effectiveness metrics:

- time required to complete the experiment and
- number of attempts to complete each task.

From these metrics, we calculated four performance indicators:

- success rate for each officer,
- average number of attempts to complete the experiment, and
- error-free performances for all officers combined.
- average time required to complete the experiment

We evaluated overall officer satisfaction from the participant questionnaires and included the open-ended officer comments to calculate the mean officer satisfaction.

Table 1 shows success rates for each police officer. Of 26 participants, 20 had a 100 percent success rate indicating that they completed all tasks. The six officers who had less than 100 percent success made a few attempts to complete each task, then quit after some failures. Average satisfaction correlated positively with the success rate (Pearson’s r = 0.478). The overall success rate was 92.31 percent.

Many officers could not complete tasks in their first attempt. Hence, they had to repeat the query to retrieve the information. The average number of attempts to complete a query was 1.738. This result represents an important cutoff because the average satisfaction rating for officers who took more than two attempts was 86 percent. In contrast, the average satisfaction rating for officers with one or two attempts was 90 percent.

Because the participants were all first-time VoiceLETS users, we expect the average number of attempts to complete a query will decrease as officers become familiar with the system in every day use. One officer specifically commented...
The usability experiments and analyses confirmed VoiceLETS’ effectiveness in offering voice-based access to the LETS database from the field. Several Alabama law enforcement officers in various counties are currently using the system. Improvements to be incorporated in the next version include an earlier beep that lets officers speak sooner and shorter voice prompts that let them complete tasks in less time. We also need to extend the VoiceLETS grammar to capture spoken queries with a southern accent.

We hope VoiceLETS offers a model for future law enforcement systems incorporating speech-recognition and audio interfaces to mission-critical information-retrieval applications.

**TABLE 1**

Success rates for 26 officer participants correlated with average satisfaction with VoiceLETS interface experience.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success rate (percent)</td>
<td>100</td>
</tr>
<tr>
<td>Number of officers</td>
<td>20</td>
</tr>
<tr>
<td>Average satisfaction (percentile)</td>
<td>90.93</td>
</tr>
</tbody>
</table>

Figure 4. Officer satisfaction levels for each post-test question. Results show consistently high satisfaction rates.

Juan E. Gilbert is an assistant professor in the Department of Computer Science and Software Engineering at Auburn University. Contact him at 107 Dunstan Hall, Auburn University, AL 36849-5347; gilbert@auburn.edu.

Richard Chapman is an associate professor in the Department of Computer Science and Software Engineering at Auburn University. Contact him at 114 Dunstan Hall, Auburn University, AL 36849-5347; chapman@eng.auburn.edu.

Sangeeta Garhyan is a graduate research assistant in the Department of Computer Science and Software Engineering at Auburn University. Contact her at 107 Dunstan Hall, Auburn University, AL 36849-5347; sangsp@yahoo.com.

**REFERENCES**

