Exclusively for ICMI Members

Call Center Management Review®

Strategies & Solutions for Call Center Professionals Worldwide September 2002

Technology Focus with Mark Levinson

Voice Technologies Offer Higher Potential for Better Service, Lower Costs

Voice applications can be designed to support overall customer relationship and branding efforts.

When Continental Airlines replaced its touchtone fight information system with one featuring speech recognition, the number of callers exiting the automated system to live agents dropped by 50 percent. Amtrak experienced similar results after releasing a new voice-powered train schedule and reservation application. And Tellme's 800-number directory assistance application completes more than half of all calls without the need for live agents.

Voice technologies – speech recognition, text-to-speech and speaker authentication – are enabling a new breed of interactive voice response (IVR) customer-contact applications. At their simplest, they streamline existing touchtone-based systems. But they also enable a wide variety of new, more complex information requests and transactions previously not practical or possible with touchtone and which, in well-designed systems, can make the caller's experience faster, easier and more satisfying than ever before.

More importantly, with the ability to really engage callers in conversation and convey mood and personality through the "sound and feel" created by the application's "voice" and audio production, these technologies create a new, unique communication medium – one that's as powerful as the Web.

Understanding the Technologies

To make the most of these technologies, it's important first to understand their capabilities and limitations. Then the applications must be designed to use them well – a task that demands some skill and experience. And finally, the fullest advantage is gained by aligning and integrating voice-powered IVR applications with other customer-contact methods, including live agents and the Web. Let's take a closer look at the three voice-based technologies:

- **Speech recognition.** Although speech recognition has been around for years, until recently, it failed to live up to its hype. But the technology has finally matured; to put it simply, it works. It's dependable even for critical applications. That's not to say that it works 100 percent of the time. It doesn't. But the current technology approaches human performance within the limited subject domains in typical applications. A well-designed system will engage callers in a dialog and ask them to repeat anything that's not recognized, just as an agent would.

- **Text-to-speech (TTS).** Text-to-speech entails the audio synthesis of spoken words from text input. Since this can be done on-the-fly as needed, its huge advantage is that it eliminates the need to record prompts during application development (which can easily number more than a hundred in a good-sized customer service application) and, as importantly, it requires no maintenance – no rerecording of prompts over the life of the application as it's updated and its content evolves.

  Until recently, most text-to-speech engines' voices sounded robotic and were perceived by many as unpleasant and hard to understand. Newer products use concatenated phonemes or syllables, where reference recordings of a voice talent speaking a variety of sentences are chopped up to create a library of individual sounds. These sounds are then reassembled as needed by the TTS engine to form the desired words.

  The result is a big improvement over previous TTS methods, but the voices, at best, are still somewhat mechanical and emotionless. One compromise approach is to have a human voice talent record prompts for the static parts of the application, and use TTS only for portions where the information being spoken back to the caller is very variable or constantly updated. However, changes of voice within a given transaction can be jarring and confusing to callers, so it's best to record the static prompts with the same or similar voice.

- **Speaker authentication.** Speaker authentication (or speaker verification) is a relatively recent addition to the scene. It uses voiceprints unique to each individual. It's not a critical technology, since its basic functionality can be duplicated

with PIN codes, but it offers several significant advantages.

The technology can allow "transparent" authentication in which callers speak a bit of routine information, like an account number, which is matched against their voiceprint, so no separate password is needed. Or it can provide increased security over a simple password or PIN code. Security systems using passwords or PIN codes perform authentication by requiring something you know. Stronger authentication is gained when the system requires both something you know and something you have. To use an ATM machine, for example, you have to have an ATM card and know your PIN number. With speaker authentication, what you have is your voice, which is always handy (vendors claim the systems work well even if you have a cold!).

One issue with speaker authentication is that, because it employs a statistical matching of input data to a reference, there will always be some level of error: false positives (where an impostor is accepted) and false negatives (where the legitimate user is rejected). The error levels are typically low – on the order of a few percent – and can be adjusted to reduce false positives at the expense of increased false negatives (for example, in applications where high security is the dominant requirement), or the reverse (in less secure applications, like voicemail, where the priority is ease of use).

Putting the Technologies to Work

So how can you make the most of the capabilities offered by these new technologies? To start, it's important to realize that applications using them are more complex and subtle than conventional IVR. They enable more complex transactions to be automated and must deal with the inherent variability of spoken interactions. Using them well is not simply a matter of sketching out a call-flow diagram and writing code.

The voice user interface is the means by which the application orchestrates its interactions with callers. A good user interface will make the application simple to understand, fast and pleasant to use. Voice user-interface design is too extensive a subject to be covered in one article. But as a general principal, a good design will not only engage callers in an easy-to-understand give-and-take, it will maintain their trust and comfort levels by explaining what it's doing and why. You have to tell callers what kinds of things the system expects them to say, and it should recover gracefully when it doesn't understand.

An equally important part of the user interface is its "sound and feel." What most affects users' perceptions of an application and, by extension, the service or company it represents, is not the fact that the system understands what they're saying, but rather the naturalness and "personality" it seems to have. A good "sound and feel" will be engaging and entertaining to customers in ways that the Web, for example, is not. With the use of professional voices, audio logos and music, it provides excellent opportunities to support your company's brand image.

The design process shouldn't stop with the handing off of a design document to the developers. It's vitally important to test the application with real customers and "tune" it as needed. Data can be obtained with focus groups on a prototype system and by recording sample calls when the system goes live. Because the conversational nature of voice systems allows a wide range of caller responses, you will almost certainly hear people say things you didn't expect. And you can get upclose and personal information about callers' reactions to the system by hearing what they say and their tone of voice.

Bringing It All Together

OK, you've got plans for great voice-powered IVR applications. How do you make sure their value to your organization is maximized? The goal isn't simply to build a good IVR system. It's to offer the best customer service while reducing costs. Voice-powered IVR systems can go a long way toward meeting these objectives, but you can go even further by aligning the customer experience of your voice applications with those of live agents and the Web, and designing them to jointly support your overall customer relationship and branding efforts. Good intermodal alignment presents customers with a contact environment that reduces effort and confusion by supporting consistent mental models of the transactions offered and how they work. Let's look at examples:

■ **Voice and live agents.** Voice applications can increase the efficiency and cost-effectiveness of call center staff by offloading high-volume transactions, thus allowing agents to concentrate on less common, complex ones. Customers will accept and use the automated system if the terminology and types of transactions offered are similar to what they would experience with live agents. With a well-designed user interface, the caller's experience can be as pleasant as dealing with an agent, and in many cases, faster and more efficient.

The automated system can also help to direct callers to the appropriate resources and save time by obtaining routine information like account numbers. But the effort will backfire if it appears to callers that the information isn't being used. For example, the application asks for an account number, which is used to route the call, but that data isn't popped to the agent's screen. The agent asks the customer for the same information. From the customer's point of view, she's just wasted her time dealing with the automated system. She'll try to bypass it in the future – and she has been left with a negative impression of the company.

■ **Voice and the Web.** The manner in which people interact with a Web site, using text and graphics, is quite different

Hearing Is Believing

Here are three well-designed, publicly accessible voice-powered applications that you can try.

- Amtrak schedules and fares
  800-USA-RAIL
  Try checking various train schedules.
- Tellme voice portal
  800-555-TELL
  News, stock quotes, traffic, sports, etc. Get your horoscope or browse local movie listings.
- United Airlines flight information
  800-824-6200
  Try getting flight arrival or departure information.

than with an IVR, which uses sound and speech. However, both should adhere to and support a common conceptual model.

You should use the same terminology and try to offer the same categories of information and the same transactions. It's important to have a single password for both the IVR and the Web. Requiring customers to have two different passwords for one company is sure to be annoying. Moreover, as previously discussed, by giving your voice applications a "sound and feel" that's consistent with the "look and feel" of your Web site, you can more strongly reinforce your desired branding messages.

These new voice technologies offer the potential to dramatically reduce costs, while providing better service to customers. By designing them well and integrating them with other customer contact methods, you can fully realize their potential to create an effective customer relationship medium. CCMReview

Mark Levinson

Mark Levinson, Principal, VoxMedia Consulting has more than 15 years' telecom industry experience, including managing design, development and deployment of real-world voice technology applications. He can be reached at mark@voxmediaconsulting.com or 781-259-0404.