

# Computing With Voice

Angela Kay Juul (E-mail: profjuul@hot.rr.com), University of Mary Hardin-Baylor  
Patrick V. Jaska (E-mail: pjaska@umhb.edu), University of Mary Hardin-Baylor

## Abstract

*Recent and continuing improvements in voice recognition and voice response technology have provided tremendous customer service benefits for users and implementers alike. Advantages for business organizations such as improved quality and productivity, reduced costs, increased return on investment, use in many industries, and increased standardization show promise for this technology now and in the future.*

*In this case study, we will discuss the procedure for choosing and implementing a successful interactive voice response (IVR) system. Included in this study is a discussion of the benefits of an interactive voice response system. These benefits included improved customer service and increased productivity without increased staffing.*

## 1. Introduction

A description of the importance of voice response and voice recognition is given by M. McGonagle (1994): “Humans communicate principally via spoken language, with writing being used primarily when the communication is not going to be immediate and interactive.” Thus, for the most natural interface, the goal is to have computers understand voice commands, process them, and speak back results in a natural-sounding voice.

In looking at the technology that is the basis of such systems, applications can run on personal computers and servers, as well as UNIX, Macintosh, and other platforms. The hardware at the core of voice processing is the soundcard, which experienced tremendous improvements and availability in the 1990’s, thus increasing the volumes of desktops that could process speech input and output (McGonagle, 1994). Improvements in this hardware have continually enabled the systems to process faster, with higher quality, providing an increasingly satisfied user.

In this paper, we will first give a technology overview of voice recognition and response systems. Next, we will then review a case study application of an interactive voice response (IVR) system. We will then present a summary of the results of the case study and procedures for implementing a successful IVR system.

## 2. Technology Overview

First we need to make a distinction between voice recognition and voice response systems. Voice recognition systems can recognize human speech. Usually the system consists of a device for the user to speak into (microphone or telephone) and a computer system with voice recognition software. Voice response systems are usually menu driven systems where the user selects different options using a touch-tone phone or minimal voice input.

### 2.1. Voice Recognition Systems

In voice recognition technology, products are normally classified as using either continuous or discrete recognition. The continuous recognition is similar to spoken conversation, whereby words flow freely, one right after the other. To contrast, discrete recognition requires a short pause between each word. This demonstrates one of the conflicts between humans and the voice recognition technology: All other things being equal, humans would prefer

to talk in their natural conversational style, yet conversational functionality is far more difficult for the computer to detect word boundaries (McGonagle, 1994).

Another distinction between types of voice recognition systems is the size of vocabulary. A voice recognition application with a small vocabulary is generally one with less than 1,000 words or phrases, and one with a large vocabulary has more than 1,000 words or phrases. By separating large systems into multiple smaller systems, each of the system components can use its own vocabulary, keeping the overall system more manageable. A smaller vocabulary means that the computer has fewer possible words to choose from, thus reducing needed computer resources and thus increasing processing performance (McGonagle, 1994).

The third type of categorization of voice recognition is that of speaker dependency. In speaker-dependent systems, the system needs to be trained to recognize each user's individual voice patterns, while speaker-independent systems use a model that can recognize any speaker's voice. This is an area that has had improvements, according to T. Weaver (personal communication, 10/16/2002). Says Weaver, "I think the technology of voice recognition has evolved. The software is quicker and the language is cleaner than in the past. Initially, the training takes a long time. Everyone has a different accent or way of pronouncing words. One of the amazing things is that it 'learns' what you mean." To contrast, Conversay's commercial application of a Distributed Speech Recognition (DSR) pilot with a handset manufacturer and wireless service provider will use a speaker-independent system, thus can serve thousands of different voices. Additionally, the DSR splits the speech recognition processing between both the handheld devices and servers, minimizing the processor and memory requirements. This speech technology is further extended to wireless device manufacturers due to recent developments in QUALCOMM'S Mobile Station Modem chipset (Conversay announces, 2002).

## **2.2. Voice Response Systems**

In voice response technology, systems are divided into digitized versus synthesized voice systems. Digitized voice systems require words or phrases to be recorded, indexed, and saved. The saved words and phrases are then pieced together to form sentences that are played to the caller upon command. Synthesized systems convert text to speech, allowing for fewer limitations on the size of the vocabulary (since the voice is not recorded). The quality in synthesized systems is limited to the hardware and software, the lower end sounding mechanical, and the higher end sounding more natural with pitch and tone more similar to human voice. Interactive Voice Response (IVR) provides a computer interface with voice. IVR provides efficient applications accessed by telephones, but also makes sense for devices with small screens, as IVR allows the delivery of information without having to display data (McGonagle, 1994).

Putting voice recognition with voice response means that a customer can dial a phone, ask a question, and receive an answer verbally from a computer. An example of this is an order entry product called BigShop, made from the German speech recognition integrator SpeechCo, who helps create solutions and applications with large grammars for industry, commerce and service applicable for information and transaction systems. The speech-enabled automatic order-line BigShop uses speaker independent automatic speech recognition and can use either recordings or speech synthesis. The system is able to gather information and ask for article features until all attributes of an order are complete, and then integrates with existing merchandise control systems to complete the order. The security in place asks for the caller's customer number and birth date, similar to that used for the Web system security (Speech-enabled order system, 2002).

Experts caution that companies should not remove the touchtone menus after voice recognition is in place, because customers are familiar with the menus and because touchtone provides a very inexpensive way of servicing the customer. Instead, business can be improved by adding voice recognition in addition to IVR, providing new levels of products and services to customers. For instance, a brokerage may see customers do more trading with the firm simply because stock prices are easily accessible via voice recognition (Hollman, 2002).

### **3. Case Study**

A mid-sized nonprofit educational organization in Texas had a reputation for offering high quality services to school districts. Among many services offered, membership in its cash management program was substantially increasing. Customer service satisfaction levels were high, however with the increased membership and proportionally increasing manually-processed transactions, strategic planning indicated that increased staff and office equipment would be necessary in the future, unless more productivity could be generated with existing staff. Backend systems were adequate and had not given any trouble, but because of technological improvements in the industry, the modern database benefits of improved stability and recoverability were desired. In short, the problems to be solved were to increase productivity while keeping staff levels unchanged, and protect the cash management system with new database benefits.

#### **3.1. The Team**

An internal team was assembled to aid in the implementation of the project. The internal project team consisted of user area management, a Business Systems Analyst/Project Leader, Telephony Manager, Network Communications staff, Database knowledge experts, and Information Systems Management. Once the selection was narrowed down to two companies, additional staff and management were included in the vendor decision.

#### **3.2. The Process**

The key to the project's success was the methodology, communication, and teamwork. The internal team began by creating a requirements analysis and high-level project plan, which directed the information systems staff to the next step of a technical vendor selection. As companies were researched and interviewed, Gartner Group and other communities were consulted to gain an understanding of the major vendors who could contribute to a solid system. Concurrently, a design was created internally, being reviewed by all members of the team, and budgets updated to include projected costs, including those costs for the new database. Once the base software company was chosen, a sizing analysis was done to estimate the number of software licenses needed. The Telephony Manager, Business Systems Analyst, and Network Communications Staff worked directly with the software provider's technical staff to install the hardware and software "agents". Once the base software was in place, application development began, and the Business Systems Analyst stayed with the project on a daily basis, reporting to the team during weekly status meetings and project reports. It was decided that the first phase of the system would be for the IVR portion only, with a second phase to include the Internet addition.

During development, an important decision was made to make sure that the voice used on the system was one from a seasoned manager who had significant probability of staying with the company for an extended number of years. In using one consistent voice, customers become comfortable with one voice, with all phrases in the same voice, now and into the future. As the voice phrases were recorded, the scripts were coded, and the phrases were then attached to the code, ready for testing. The project team employed the help of Quality Assurance for testing, and included additional test helper employees to ensure the system was easy to use according to the user documentation written. Stress testing was done, to evaluate what happened when all "agents" were used at the same time, both from what the user heard on the phone to how the system processed all the transactions during the stress time. After thoroughly checking test results, the project continued with the next step of a thorough system audit. A well-respected third party EDP auditor was brought in to both audit the system and evaluate its integrity. After rigorous reviews, the audit reflected a superior, stable system.

A target go live date was solidified, project plans were updated, and staff training for the information systems staff was coordinated and executed. The customer service representatives were thoroughly trained on the system, and their roles were changed to be used on more of an exception basis, only if the customer opted out of the IVR menu to speak to a live person. The customers were then notified about the upcoming change to the new system approximately 30 days in advance.

Production data was loaded and the system went live, smoothly. At first, there were many callers who opted to speak to a person. Once the customer's questions about the system were answered, more and more chose to utilize the system. Additionally, customers could get information about their accounts after normal business hours, which was a new benefit. Within weeks, many customers chose to use the automated "agents".

Once the system was in production for 30 days, the additional pieces necessary to implement the Internet functionality were put in place. Although the installed program code could be used for both the IVR and Internet, there were several Internet communications aspects of the project that needed to be done, and additional coding for the extra features the Internet site would have above and beyond that of the IVR. High level encryption was used, and agreements with browser software vendors had to be gained in order to provide high-level encryption software to customers. User instructions and team training were provided so that customer installation of high level encryption would be easily accomplished. Additional information such as daily updated current day news highlights were placed on the Internet site. Decisions were made reflecting the style and color of graphics to be used on the page. Customer statements and reports were added, as well as printable transaction confirmations.

Internal management reports were designed, created, and routinely generated for user management to demonstrate and summarize the number of calls and transactions taken, including incremental numbers by hour.

### **3.3. Vendor Selection**

Vendor selection occurred in two phases: One vendor to supply the base software, and another to actually provide the scripting. The company utilized for hardware and software was chosen due to several factors. Edify provided software "agents" that could answer the phone, ask for passwords, offer information to customers, and process transactions. Additionally, the technology worked on both IVR as well as Internet platforms, meaning that one program could be written, and the system could be offered both over the telephone and the Internet. The secure way in which the system sat on the Internet firewall was of primary importance, providing the level of safekeeping mandatory to maintain high security standards. Multiple customer references were pleased with Edify's service and system provided, and telephony user groups agreed this was a vendor who could both deliver now and provide vision into the future.

The method of implementation was then evaluated, whether to code internally or outsource the work. The option of internal staff training and subsequent system development was considered, as well as the potential utilization of Edify's optional Professional Services. The internal staff option would have a long learning curve, would need multiple people to be educated due to turnover possibilities, and would potentially need additional consulting services for the first application. The technology was viewed as a highly specialized area that was not fundamental to the organization's core business. In evaluating Edify's Professional Services, the cost was higher than desired and included travel between California and Texas. These financial factors warranted more research to find an application developer that provided high quality and cost effectiveness.

With Edify's help, the team located a company who had a proven track record of implementing and programming Edify's systems with excellent results in a cost-effective manner, and included customer accolades noting superior service. Vexis, formerly known as VoiceXpress, headquartered in Tulsa, Oklahoma, was the professional services company selected for application development. Vexis verified that the hardware and software "agents" were installed correctly. They fine-tuned all aspects of the system for optimal performance. Vexis provided the application development for the design that the team had modeled internally, and gave guidance for the voice scripts to be recorded. They attached the application to the new MS SQL database, then tested, implemented, and subsequently maintained the system.

### **3.4. The Results**

The number of customer members continued to grow, and additional staff was not needed to process the increased number of transactions. By utilizing the IVR application to perform the majority of the transactions,


processes improved so that existing staff worked primarily on the more complex questions requiring human intervention, on an exception basis.

For lessons learned, the team learned that the end Application Development Provider (in this case, Vexis), could have provided the Edify Hardware and Software “agents” in addition to the scripting services. The benefit of this business practice is that there is only one source needed to call when a question arises, instead of having to discern the nature of the inquiry prior to inquisition.

As of this day, the system continues to provide the levels of security, availability, reliability, and recoverability mandatory for a cash management system.

#### **4. Summary**

In this case study, a procedure for choosing and implementing a successful interactive voice response system was given. The successful implementation of this system included: (1.) choosing the appropriate internal project team to analyze the needs and requirements for the organization, (2.) considering and adopting a customer focused approach to determining the requirements of the system, (3.) a well-defined vendor selection procedure, (4.) employing quality assurance procedures throughout the project, (5.) a thorough EDP audit of the system, (6.) thorough system training for customer service representatives, and (7.) an internal management reporting system to monitor the performance of the system. Following the above procedure resulted in the successful implementation of a reliable, customer focused IVR system.

In this paper we reviewed some of the aspects in which voice recognition and voice response technologies are improving the way people live, learn and work. Voice technology positively affects the way people receive and access information for personal and business use. The business bottom line efficiencies are proven powerful now and into the future with resulting cost reduction, accident reduction, profit increase, and better quality delivered to a more satisfied end customer. With all of these attributes in favor of voice technology, the best results are most likely yet to be seen. 

#### **References**

1. Conversay (2001). “Selecting a speech server platform: Key technical and financial considerations” (Version 1.0). Retrieved October 24, 2002 from <http://www.conversay.com/shared/pdf/Selecting%20Server%20Platform.pdf>.
2. “Conversay announces commercial deployment of distributed speech recognition solution” (2002, October 15). *SpeechTECHNOLOGY Magazine*. Retrieved October 24, 2002, from AmComm Holdings LLC Web Site: <http://www.speechtechmag.com>.
3. Hollman, L. (2002, November). “Let your voice do the walking”, *CallCenter Technologies, Services and Strategies for Contact Centers*, 15, 28-40.
4. McGonagle, M. (1994). “Voice-recognition and voice-response systems”, In J. Keyes (Ed.), *The McGraw-Hill multimedia handbook* (1st ed., pp. 35.1-35.27). : McGraw-Hill.
5. “Speech-enabled order system BigShop discharges call centers” (2002, October 15). *SpeechTECHNOLOGY Magazine*. Retrieved October 24, 2002, from AmComm Holdings LLC Web Site: <http://www.speechtechmag.com>.
6. Yoshiko, H. (2002). “Multilingual voice recognition middleware targets automotive applications”, *Electronic Engineering Times*, (1203), 24-26. Retrieved October 16, 2002 from Academic Source Premier (0192-1541).

Notes