What Is Speech

RECOGNITION

sas
Specialty Answering Service

WWW.SPECIALTYANSWERINGSERVICE.NET

Contents

1 Introduction to Speech Recognition

Today, speech recognition technology has advanced to such a stage that you can have a real conversation with a computer where it not only understands what you say, but is also able to communicate back to you.

For a long time, speech recognition was in the research stages where practical uses were minimal. But, around the late 90s, the technology grew at a rapid pace and it became a financially viable option for several industries, especially call centers, which started using speech recognition in combination with IVR systems to not only enhance the customer experience, but also to cut costs.

Some of the key factors that contributed to the rise of speech recognition include computers' faster processing capacities and their drop in prices, as well as advances in speech recognition algorithms, which improved their accuracy significantly.

In the early days of speech recognition, it was thought of as a technology with a great amount of practical potential. Research predicted that speech recognition would be a $1 billion industry by the end of the 90s. It was expected that in North America alone, call center application use of speech recognition would grow from $60 million to $150 million in the period 1996-2001.

Today, most large call centers are actively considering speech recognition as part of their technology strategy, although negative customer perceptions and low real life accuracy rates have led to some bad press for speech recognition in recent years.

There are several products and services that fall within the gamut of speech recognition. The underlying technology of speech recognition is the speech recognition engine – which converts the spoken word into the written word. Related products in this field include text to speech engines, voice platforms, development toolkits that help software professionals to build, maintain, and fine tune speech recognition applications, as well as tools to monitor the accuracy of those applications. Today, there are 'plug and play' packages that allow call centers to customize and implement the solution within a matter of a few months.

In the early days of speech recognition technology, the focus was on fine-tuning the speech recognition engine with different kinds of algorithms in an attempt to improve the accuracy of the engine. However, this has been successfully conquered, and application providers are focusing on providing value added features such as ease of use to enable faster acceptance in the industry.

2 Evolution of Speech Recognition Algorithms

Speech recognition has come a long way in terms of understanding normal conversations. This has drastically cut down the error rate as algorithmic improvements in the last decade have managed to bring down errors by about 30% every year. This has made speech recognition a viable practical solution.

Leading speech applications can now achieve accuracy in the range of 90%-98% depending on the nature of the transaction. One of the key reasons why accuracy has improved is because the vocabulary of most current applications includes more than 30,000 words as compared to a few dozen words in the early 90s. This in turn enables developers to incorporate the ability to respond to complex customer requests that can be phrased in more than one way. Contextual understanding has also contributed to improving the user experience while using these applications.

Another key algorithmic change is the ability to correctly interpret 'natural language' or continuous speech. In the initial days, users had to speak specific phrases in clipped form for the engine to recognize and respond effectively. However, the latter versions of the software allow people to speak in full sentences in a natural manner, thus making the interaction more real and life like. This also helps to drastically reduce call time by allowing users to request multiple services simultaneously. Practically, call centers have reported as much as 30% to 40% reduction in call time due to this feature.

Even after the advent of speech recognition technology, for a long time call centers remained content with just IVR systems. But common standards such as VoiceXM and SALT, as well as support from established vendors such as IBM and Microsoft, made sure that applications based on speech recognition technology were adopted by the mainstream. Some of the key cost savings arose from increased automation of calls, partial automation of calls routed to agents, and more accurate skills-based routing of calls. Indirect benefits also included increased customer satisfaction, better brand awareness, and the ability to retain customers who would otherwise switch to competition with this cutting-edge technology.

Initially, speech applications were deployed in industries such as financial services and airlines where it was difficult to enter ticket symbols and travel dates onto a keypad. The huge success of these implementations resulted in other industries adopting the technology. However, in the early days, speech systems only replicated the IVR menu structure with prompts such as "Press or say 2". Thus, instead of flattening the menu structure, it resulted in longer calls due to the long prompts. Soon, the industry realized that in order to fully benefit from the technology, a well-designed voice user interface (VUI) was necessary. A well-designed VUI had to cater to a first time caller as well as an advanced user who may want to skip several steps forward or get multiple information requests processed simultaneously.

Nowadays, sophisticated applications employ statistical language modeling and open grammar algorithms to statistically infer the reason for the call. This technique helps in effective call routing,

by asking callers open ended questions such as “Please explain the reason for your call”. These algorithms rely on a statistical analysis of a user’s input rather than a one-to-one match of every input to a pre-determined set of responses. This enables the system to understand a much larger set of inputs and provide a probable interpretation and response.

A well-designed VUI should have a personality and tone that is pleasing to the caller as callers often fail to distinguish between a VUI and a live agent. The first step in designing a robust VUI is to evaluate inbound calls to understand the way customers interact with live agents and the way they phrase their questions and responses. After this, various prototypes are designed and tested prior to choosing the right one to implement. The key thing to remember is that designing a good VUI is an ongoing process that requires initial planning as well as ongoing monitoring and refinement. This ensures that an optimal system is established to meet the callers’ needs while avoiding frustrations caused by incorrect responses from the system. In turn, this will help the application to reach its maximum potential, providing a good return on investment.

3 Practical Applications of Speech Recognition Technology

Some of the key industries where speech recognition finds practical applications include banks and financial institutions, airlines and telecommunications. Telecom companies use speech recognition for offering services such as directory assistance and to complete collect calls. Airlines use it for flight reservation by passengers as well as flight scheduling by employees. Logistics companies use it for tracking packages in transit. In financial institutions, speech recognition has been used for handling balance inquiries, responding to customer requests, helping customers execute simple transactions such as applying for mortgages, getting stock quotes, and so on. Some of the key financial institutions like American Express, Charles Schwab & Co., and NationsBank have successfully used speech recognition technology to better serve their customers.

Studies have shown that callers are less likely to ask for a live agent when catered to by a speech application versus a touch-tone menu. Two key factors contributing to the wide acceptance of speech applications by the callers are improved voice interfaces and the use of cell phones, which makes touch tone responses cumbersome to handle while on the move.

One of the key advantages of using speech recognition technology is that when used in the right manner can help to cut operational costs by a large margin. For example, a human agent would cost nearly $30,000 per year whereas voice recognition can cost around $3,000 per phone line. Thus, the return on investment can be seen in as fast as a few months.

Although brokerage firms were among the early adopters of the technology, banks were not too keen on introducing the system. This was due to integration issues with legacy systems, and the multiple applications that were typically used to handle different products and sales channels offered by the bank.

One of the key success factors of a practical application is that it is modular and can be integrated with other software used in the call center. This is because call centers often want to implement the solution in phases, and unless the application can talk to the databases and other applications, it will not be widely used.

3.1 Real Life Examples

In this section, we will look at some of the real life examples, where successful implementation of speech applications has provided benefits for call centers.

One of the earliest commercial deployments of speech recognition technology was the VoiceBroker application implemented at Charles Schwab & Co in 1996. It was used to provide customers with stock quotes and information about other market indicators. Thanks to the efforts of the company to analyze each error and constantly tweak the system, the error rate has dropped from nearly 30% to about 4%-5%. Schwab has also implemented biometric speech software to verify a caller's identity with the help of his or her voice. This is done by matching the voice against a stored

electronic voiceprint. The financial service industry is quite apt for speech recognition technology use as it receives a large volume of calls every year, and even a marginal reduction in call duration can provide huge cost savings.

Fidelity Investments has used a natural language application called FAST to help users perform transactions such as obtain stock, option, mutual fund and index quotes, review account balances, and even execute trades. In fact, more than three quarters of calls that reach the Fidelity call center are handled automatically through the IVR and speech recognition platforms, thus freeing up agent time to focus on sales and efficiency improvements.

Standard Life uses speech recognition for its life insurance and pension businesses, making it possible to authenticate callers, identify their requirements, and properly route calls. It also relays caller information to the agent, so that callers do not have to repeat the information again. This has helped them increase call-handling capacity by 25% and call routing accuracy by 66%. At Standard Life, the VUI is personalized and known as 'Sheila'. The voice application has not only improved call routing, but also increased customer satisfaction levels.

The general insurance company Suncorp has used Natural Language Speech Recognition (NLSR) to replace its IVR system. The grammar of the application includes over 0.1 million phrases which are specific to the financial services industry. The company has managed to reduce its call wait times to about half a minute, and misdirected calls are virtually non-existent.

Ladbrokes, a betting giant, has managed to successfully divert calls based on their nature. Transactions such as placing a bet or an inquiry of odds were handled automatically while requests for more 'customized' bets were routed to agents. The company has a large database of horses and football players that are updated on a real time basis and feeds the data into the voice application.

4 Advantages of Using Speech Recognition in Call Centers

The initial investment required for a practical application of the software is considerably large and it can be to the tune of 0.5 – 1 million. However, for a large call center that handles several million calls per year, even a savings of 10 cents per call can be a considerably large amount. The biggest cost savings comes from automating a larger number of calls, thus reducing the need for live agents.

4.1 Cost Savings Calculation

A sample calculation for expected cost savings and revenue enhancements that a call center can generate by using voice recognition software is given below:

Assumptions:

- No. of incoming calls per year: 3.6 million
- % of calls that go to a live agent: 34% of 3.6 million = 1224000
- % of live agent calls handled by a speech recognition application: 65% of 1224000 = 795600
- Average duration per call: 3 minutes
- Average cost per live agent: $30,000 per year
- Telecommunication line charges: $0.10 per minute

Before the introduction of speech recognition application:

- No. of calls that a live agent can handle in a day (360 minutes of call time per day) : 360/3 = 120 calls
- No. of calls per year per agent : 120 * 22 * 12 = 31680
- Total live agents required: 1224000/31680 = 39 agents

Before the introduction of speech recognition application:

- Savings in call duration due to speech recognition application: 20%
- Average duration of calls which pass through speech recognition engine: = 3*80% = 2.4 minutes
- No. of calls in which time saving is obtained = 795600

- No. of such calls that a live agent can handle in a day (360 minutes of call time per day) : 360/2.4 = 150 calls
- No. of calls per year per agent : 150 * 22 * 12 = 39600
- Total live agents required for such calls: 795600/39600 = 20 agents
- Remaining calls = 122400 – 795600 = 428400 calls
- Total live agents needed for remaining calls = 428400 / 31680 = 14 agents
- Total agents: = 20+14 = 34 agents
- No. of agents saved = 39-34 = 5
- Savings in personnel costs: 5* $30,000 = $150,000
- Savings in telecommunication costs: = 477360 * $0.1 = $47,736

Similarly, if you assume retail revenue of $600 per account, even a 1% increase in sales achieved from the increase in sales time available will work out to be a large amount.

One of the chief advantages of speech recognition applications is the fact that they take away the cumbersome nature of IVR, making it unsuitable for carrying out complex transactions without the help of a live agent. In the initial days of its success, callers also liked the fact that it offered the ability to obtain information and conduct complex transactions like those that you would do over the Internet or through a phone call. However, with the widespread popularity of smartphones with built-in internet browsing capabilities and the spread of 3G and 4G, the advantage of speech recognition is slowly ebbing away. This is because customers today conduct transactions either through the web or through mobile commerce rather than waiting for the speech recognition engine in the call center to understand and respond to their requests.

What speech recognition has done is to make the IVR system in a call center more user friendly by allowing the customer to do more transactions in a faster and more natural manner.

4.2 Getting Ready for Speech Applications

If you are considering implementing speech recognition systems in your call center, then answering some of these questions may help you reach a decision:

1. Is your average wait completion time for simple transactions such as routine balance inquiry too high?

2. Do agents spend too much time in non-value adding activities such as validating a caller's identity, routing the call to the correct department, and answering simple questions?

3. Do your callers call while on the move, making it difficult to navigate through a touch-tone IVR menu?

4. Are you operating in a highly competitive industry where switching costs for your callers is less than your customer acquisition costs?

If you answered any of these questions in the affirmative, then evaluating speech recognition applications as a means to enhance service levels and productivity in your call center would be of benefit.

4.3 Cost-Benefit Analysis

Before implementing speech applications, the call center should perform a cost-benefit analysis to choose between a complete redesign of the call center versus a tactical piecemeal add-on approach. A high ROI application that is easy to implement would help to get the necessary buy-ins from all stakeholders for future rollouts of more applications. It is important to have a complete plan in place so that the infrastructure and vendors chosen for the initial phases are strong enough to handle subsequent phases of the project.

Evaluate the project not just from an ROI perspective, but also from the point of whether it is in alignment with your organizational objectives, and whether it would help you achieve your strategic, financial, marketing, and customer service goals. Some of the other questions that you should answer, before you go ahead with a speech application are given below:

1. Is your overriding focus on cutting costs or on improving customer service? While most call centers would want a mixture of both, it is always better to identify which is more important than the other, in the event that you need to make some trade-offs.

2. Are callers satisfied with the current level of service offered?

3. Do management teams and agents feel comfortable about switching applications?

4. Who will be elected in the organization to champion the associated change management?

5. Do the speech applications give you a competitive edge over other players in the industry?

6. Do you have sufficient budgets to pay for the application and its maintenance?

4.4 In-house or Outsourced?

Once you have decided to go ahead with speech applications, the next step is to decide whether you would want to run it in-house, or would have it outsourced to a third party. Some of the questions to consider in this regard are given below:

1. Do you have the in-house expertise to customize the application for grammar, dialogue design and other human factors?

2. Do you have enough bandwidth to manage the application 24/7?

3. Do you have multi-site backup systems in place? If not, are you willing to invest in one?

Each application is different, but a system that can speech-enable most of your calls will cost you anywhere between $1.5 million to $4 million. If you do not have that kind of budgets, then it is best to contract out the project to an outsource provider who has a demonstrated dialogue design and technical expertise in this area. By outsourcing your speech application rollout, you can start reaping the associated benefits much faster than if you handle it in-house. Additionally, you will not have to worry about software upgrades and maintenance, enabling you to focus your efforts on improving the call center's core operations.

4.5 Structured Approach

If you plan to go ahead with the implementation of a speech application in your call center, then following a structured approach will improve the chances of your project's success. The guidelines below can assist with the project-planning phase:

1. *Form a steering committee* with representation from user departments such as marketing and customer service as well as departments such as the call center operations team and the technical team.

2. *Analyze call data* by collecting details on call volumes, and call types from your IVR and call routing systems. You should be able to identify specific points where the caller has problems with the IVR system and opts for the live agent. You must also analyze agent-handled calls in detail to identify specific calls that are ideal candidates for automation.

3. *Develop the business case for the project* by doing a cost benefit analysis. Compare the investments needed on the project against the cost savings and revenue enhancements that you expect. While factoring the costs of the system, you must also include the cost of integrating the speech application with the backend customer database, in order to be able to derive the maximum benefits. Include intangible benefits such as better customer service and a lowering of call drop rates.

4. *Develop a project plan* to roll out the speech application in phases. Identify the calls that are easiest to automate such as address changes and balance inquiries for the initial phase. Transactions that are more complex can be set aside for a later phase.

4.6 Evaluating Service Providers

One of the key areas of the planning process is to identify the right vendor. In this section, we will look at some of the things it is helpful to consider while evaluating potential vendors.

1. *Technical Expertise*: Ask your potential vendor for a 'dialogue specification' document and the profile of the people who would be working on your project to judge the dialogue design experience of your vendor. Ask for specific experience in the area of CTI integration for call centers as well.

2. *Data Analysis*: Ask your vendor how they plan to collect and analyze call data, as this is an integral part of designing a robust system.

3. *Implementation and Training:* Ask your vendor to share the implementation plan with you. Also, ask about what training will be offered to your staff and the costs associated with it.

4. *References:* Request references of previous implementations so that you can talk to other people who have gone through the implementation and learn about the challenges that they faced in the process.

4.7 New Opportunities for Speech Applications

Although using speech applications to further automate existing services can yield significant cost savings, today vendors are designing new applications that can automate complex transactions and provide better operational efficiencies. The transactions that best lend themselves to speech automation are those that have high volumes and are highly repetitive in nature. These include address change requests, requests for product information, and requests for switching of tariff plans. Caller identification is another area where speech automation can be used. It can enroll customers with a voiceprint to enable an added layer of security, as well as enhance customer service by avoiding the need for callers to repeat hard to remember account data every time a call is made. Yet another area where speech applications can be used is to conduct customer satisfaction surveys following customers' calls.

Some of the key advantages of implementing speech applications are given below:

- Reduce customer wait time by as much as 35%Better customer segmentation and consequent routing of calls, resulting in higher first-call resolution rates

- Higher call completion rates

- Ability to handle higher call volumes
- Enhanced customer satisfaction
- Ability to offer 24/7 service

Integrating the speech application with customer relationship management software has several benefits. The speech recognition system can be configured to greet callers by their name and communicate in the language of their choice. In addition, the speech application can also keep track of the caller's preferences in terms of preferred products and services, nature and frequency of transactions.

Recent surveys indicate that most callers enjoy being greeted by name and are happy that the system could remember their language preference and the last transaction details.

Successful deployment of a speech recognition application has always been a challenge as the majority of the work involves 'tuning' the application for your specific caller profiles. This includes upgrading the vocabulary of the speech engine with key words and phrases most commonly used by your callers. This is not a one-off process and must be repeated frequently to maintain the application's level of accuracy. In turn, this means that you would be spending time and effort on an ongoing basis with the consultants who design and install your speech application. However, you can do several things to reduce the time, effort, and costs involved in deployment, some of which are listed below:

- Implementation of packaged applications
- Expanded grammar sets including industry specific jargon
- Improving ease of use of design and tuning tools. For example, Microsoft has adapted its Visual Studio Web development tool to be used with its Speech Server, thus making it easy for developers with Visual Studio expertise to design and tune speech applications.

In fact, today's software is able to perform many functions that previously required expensive hardware. Speech recognition applications deployed over a network not only enable more automation, but also reduce the hardware required, and even support business continuity by automatically distributing licenses and workload from a failed server to a backup server.

4.8 Identifying the Vendor

Speech recognition technology can be used to reduce costs and enhance revenue. While some companies use speech applications entirely for upsell attempts, others use the technology to gauge a caller's interest for a new product or service before routing the call to a live agent for sales closure.

Some vendors also offer value-added services such as business analytics to mine automatic call details for business intelligence that can lead to increased revenues.

The most popular vendors who offer software, hardware, and consulting services for implementing speech recognition in call centers are listed below in alphabetical order:

1. Aspect Software
2. Avaya
3. Edify
4. Empirix
5. Enterprise Integration Group (EIG)
6. Genesys
7. IBM
8. Intervoice
9. LumenVox
10. Microsoft
11. Nuance
12. Sterling Audits
13. Syntellect
14. TuVox
15. Unveil Technologies
16. Voxify

In order to select the right speech application for your needs, keep the following aspects in mind:

Technology Standards: Proprietary versus Standard

One of the key changes that the speech recognition industry has witnessed in recent years is a move towards open standards and convergence with the other IT assets in the call center. Unlike IVR technologies, which largely existed in silos and made integration with backend databases difficult, speech applications are designed to enable easy communication with the other IT applications in the call center. It is important therefore to select language and platform-independent applications

so that your current IT equipment and infrastructure need not be changed to accommodate the application.

Development Options: Packaged Versus Custom

In the initial ages of speech recognition application, each application was custom made for the specific requirements of a call center. This resulted in high cost solutions that needed a long implementation time and significant effort for maintenance. Today, there are tried and true packaged applications that can be easily customized to result in low risk, low cost and quick deployments. Thus, unless you are in a niche industry with a large set of jargon and industry specific needs, it is always better to start from a packaged application.

Deployment Options: Hosted Versus Customer Premises Equipment (CPE)

If you do not have a strong IT team or you have budgetary constraints, then a hosted solution could be a good alternative. However, CPE solutions will offer you more control as they can accommodate higher call volumes when your business expands.

Vendor Expertise

Before you finalize the vendor, complete your due diligence and choose a vendor with a strong record of accomplishment that can offer you customer references. You must ideally select a vendor that can give you the software, equipment, and consultancy services for your project. This means that the vendor should be able to do the feasibility and scoping phases of your project as well as the implementation phase where you will be offered assistance in the areas of VUI design, tuning, and training. Be sure to request for post implementation support in the form of performance audits and ongoing tuning support.

Flexibility

Ultimately, the application should also offer enough flexibility to be tweaked to cater to changes in your business environment and customer profile. Thus, the system should allow the call center personnel to easily make changes through a user interface and system admin console.

The main objectives of using speech application software are to offer a service that callers are happy about, that is able to achieve measurable results, and offers a good return on investment. Most call centers are now able to realize an ROI from speech technology within the first year of implementation. However, to reap the maximum benefits, you must go through the stages of discovery, design, implementation, and optimization.

1. *Discovery:* In this stage, you would prepare a business case for speech applications in your call center. The first step is to define your business, your caller, the IT system that you currently have, and the ways in which speech applications can improve each of these.

2. *Design:* In this phase, you will design the user interface for the speech application. Unless it is easy to use, callers will not use your speech application for self-service. Make sure to reinforce your company's brand through your VUI. You can even personalize your VUI with an agent profile and a name.

3. *Implementation:* In this phase, the actual speech application is customized, tested, and deployed in your call center.

4. *Optimization:* This is a key phase that is often overlooked by call centers. In order to keep your application in tune with the changing caller behavior as well as the changing nature of your business, ongoing optimization is essential.

5 Road Ahead for Speech Recognition as a Viable Option for Customer Service

In recent times, call centers have really given a thumbs-down to speech applications. According to an IVR Survey carried out by the Call Centre Helper website in February 2012, only an average of 10% of call centers use speech recognition. As expected, the larger the call center, the higher the adoption of speech recognition. Only 6% of smaller call centers use speech applications, whereas the figure is 7% for medium size call centers and 23% for large call centers. The adoption of speech recognition has also varied widely between industries. A survey conducted in 2009 revealed that while telecom and utilities industries have almost 33% adoption rates, others such as retail and IT have not opted for speech recognition at all. The average adoption rate hovered around 8% for speech recognition in 2009 versus 10% in 2012. This shows that there has been a shift from Touch Tone IVR to voice recognition, although at a very small pace.

The main problem with speech applications is its accuracy, especially when it comes to correctly interpreting regional accents. Most systems which claim to offer as high as 90% accuracy perform much below expectations in real life and offer only up to 70% accuracy. However, survey respondents also agree on the importance of careful tuning to improve accuracy rates, which is found to improve up to 90% if tuned properly. The fact remains that there is a wide gap between recognition rates for a limited vocabulary in a controlled environment and those obtained in real life situations with a lot of background noise, vocabulary variations and accent variations. Words that are not part of the speech engine's vocabulary, known as 'out of grammar' errors, contribute to as much as 15% of the problems.

That being said, this does not signal the end of speech recognition technology being used in call centers. Its implementation will undergo changes in the future. Some potential changes would be in the following areas:

1. *Playback Information:* Customers who want fast access to information and are in a hurry do not always want to speak to a live agent. They would rather have a shorter wait time and get the information that they want more rapidly. This has been successfully used in the Dublin Airport, where a 30% increase in passenger numbers has been handled without the need for additional agents. The incoming calls are filtered according to the requirements, and calls that require info on 'departures' or arrivals' are directed to a speech recognition system, handling the requirements effectively without having to transfer the call to a live agent. The system has been fine-tuned to recognize the Irish accent, which has improved accuracy rates. The average call time for these calls is just 53 seconds, thus freeing up agent time for more complex calls.

2. *Call Routing:* Callers no longer want to speak to a person, but want to speak to the 'right person'. This is where speech recognition applications can play a major role by allowing callers to 'say' what they want so that they can be routed to the right operator, thereby leading to improved first call resolution rates.

3. *Caller Authentication:* With increased instances of identity fraud being reported from call centers, voice biometrics may provide the answer for caller authentication without having to ask for personal data. Creating a voiceprint is relatively easy and it takes less than two minutes on average to create one. This is then stored against the customer's record and the next time he calls, the voice is matched against the voiceprint stored in the database and directly routed to an agent. This feature cuts down call time spent on detailed ID checks using passwords, address, account details, and so on.

4. *Replacing Complicated IVR Menus:* Speech recognition applications will find uses in the area of 'intelligent call steering' by asking customers what they want instead of asking them to navigate through a complicated 'push button' menu.

5. *Dealing with Spikes in Call Volumes:* In industries such as betting where most calls occur just before a race, speech recognition applications can help in handling sudden spurts in call volumes.

6. *Customer Surveys:* After an agent has handled a call, a speech application can be used to conduct customer satisfaction surveys. This makes it more comfortable for the caller especially in instances where the agent is receiving negative feedback. In addition, it will also free up expensive agent time.

Some of the key reasons why speech applications have not been successful in certain cases are listed below:

- *Difficult to Use:* If callers are unable to complete the transaction through a voice application, then they will surely opt for a live agent. Another factor is the length of the call. If completing the transaction takes longer over the voice application than through the agent, the caller will prefer an agent to the application.

- *Bad Design:* A badly designed VUI is one of the main reasons why callers may opt out. Badly designed menu structures were the biggest problem with IVR systems, and this is true for voice systems as well. Instead of having multiple layers, the application must be configured to ask open-ended questions such as "What would you like to do today?" and be able to interpret all possible answers. You may also include supplementary questions to improve the interpretation accuracy of your speech engine. For example, instead of asking for just the postal code, the speech system can be configured to ask for both postal code as well as the street name so that the system is able to accurately interpret the address by getting a unique match in a larger set of situations.

- *Lack of Caller Prompts:* Callers must be prompted to provide information in a consistent manner. For example, if you have to capture a nil value, then the caller may say "nil", "no", "nothing", "zero" and so on. In such a case, instead of training the speech engine to interpret all possibilities, it is better to train the caller with a prompt such as, "If the amount is 0, please say zero".

In order to improve the customer experience, the following aspects should be incorporated in your speech application:

1. *Have help messages at the start of the call:* Callers should have an option to request the help menu at any time during the call by saying, "help". You can have a prompt at the beginning of the call that says, "At any point in time, please say, 'Help', if you need more information".

2. *Ensure that the prompts are easy to understand:* Instead of saying, "State your account number", say, "State your account number that is in the upper right corner of your bank statement. If you do not have your account number, state your telephone number".

3. *Have sufficient time lag:* Give users ample time to respond to queries and provide the information that you are asking for, before replaying the prompt.

4. *Avoid endless loops:* Placing callers in an endless loop of having to repeat a request or information multiple times is bothersome. If your speech engine fails to recognize the input after two tries, transfer the call to a live agent. Make sure that all the information captured during the call is available to the live agent and callers are not required to repeat everything from the start. You can always analyze the call later to tweak your speech engine.

6 Conclusion

While self-service applications aren't always the right choice for every business or in every situation, advanced speech applications can help companies achieve the delicate balance between giving customers a satisfying experience while successfully managing costs. Ease of use, therefore, becomes the most valuable way of differentiating your business from competitors. Voice technology has reached the stage where only small improvements in system speed can be made. Yet there is plenty of scope for improving the user experience – however complex or mission critical the requirement.

7 References

1. "Speech recognition enhances call-center capabilities", 2000, *Communications News,* vol. 37, no. 11, pp. 44-44.

2. Barbetta, F. 1996, "Unisys deliver speech recognition for call centers", *Business Communications Review,* vol. 26, no. 12, pp. 60-60.

3. Maselli, J. 2001, "Speech recognition cuts Amtrak's call-center costs", *InformationWeek,* , no. 866, pp. 24-24.

4. Zager, M. 2005, "Speech Recognition Goes Mainstream", *Call Center Magazine,* vol. 18, no. 10, pp. 20-24,26,28-30.

5. Sherter, A. 1999, "Speech recognition speaks volumes", *Bank Technology News,* vol. 12, no. 8, pp. 1-1,33+.

6. *VoiceGenie and BBN Technologies to Market Speech-Recognition Product for Large Business Call Centers* 2002, , New York, United States, New York.

7. *ALTech and Amarex Partner To Deliver Speech Recognition For Network-Based Call Centers* 1998, , New York, United States, New York.

8. "Speech Recognition", 1999, *Communications Today,* vol. 5, no. 145, pp. 1-1.

9. Kolor, Joanna. (Oct 1997). "Speech recognition gets serious" Bank Technology News, 1,43+.

10. Call Centers: Speech Recognition Centers Reduce Call Time by 35% PR Newswire [New York] 16 Aug 2005: 1.

11. Pollock, B. & Calabro, T. 2002, "Ready for speech recognition?", *Public Utilities Fortnightly,* vol. 1, no. 1, pp. 44-44.