

VOXIFY WHITE PAPER

Before You Buy Speech Self-Service: 2 Critical Questions You Need to Ask



Tell a Friend.



1151 Marina Village Parkway

Alameda, CA 94501

(P) 510.545.5000

www.voxify.com

Contents

Page 3	Executive Summary: New Speech Applications Deliver Customer Service and Reduce Costs
Page 3	Before You Buy: 2 Critical Questions You Need to Ask #1 What is the Speech Application Performance? #2 What is the Speech Application Efficiency? The Speech Application You Want to Buy
Page 4	The Challenge of Conversation Custom-Development Approach to the Conversation Challenge: Build All Call Paths SLM: A Variant of the Custom-Development Approach Toolkit Approach to the Conversation Challenge: Fixed Call Paths Traditional Approaches Cannot Give Performance and Efficiency The New Approach: Navigation Skills to Traverse Call Paths Dynamically
Page 7	Measuring Performance & Efficiency Before You Buy Measuring Performance Broad Spectrum of Call Types High Customer Adoption Across All Call Types Measuring Efficiency Deploy in Weeks Changes in Day Putting the Data to Work: Three Reference Phone Calls
Page 9	Summary

Executive Summary:

New Speech Applications Deliver Customer Service and Reduce Costs

Excellent customer service continues to be a key business differentiator. The challenge businesses and their call centers have historically faced, however, is that increasing service levels involves increasing costs. In addition, as businesses strive to further differentiate, they seek to introduce new call center services. These new services – such as a service that notifies customers when a flight is delayed or that handles inquiries about frequent flyer account details – often don't produce revenues, and as such call centers are pressured to introduce these services without increasing costs.

Enter the promise of speech applications. Speech applications came on the market in the late 1990's with the promise of liberating call centers from the age-old dilemma of trading off service levels vs. costs. The promise of speech automation was to improve both service levels and costs simultaneously. And while the applications have been improved over the years, traditional speech applications never delivered on this promise. Because of their very approach, these applications came with either cost or performance limitations. They either delivered performance – thus improving customer service levels, or they delivered efficiency – thus improving costs. But never both. Where businesses gained in one area, they lost in the other.

As such, speech applications have historically been deployed as point solutions. That is, they have been deployed to serve a narrow function because cost or performance prevented businesses from deploying them across a broad range of functions. According to Gartner, “Call centers around the world are using speech recognition solutions to automate customer interactions. Most of these are point solutions.”¹ The real value of

speech automation, however, cannot be gained with point solutions. The real value comes when an application can be deployed strategically in the call center – servicing the full spectrum of call types with performance and efficiency. Applied strategically, speech has the potential to materially impact the business. It has the potential to deliver significantly higher customer acquisition and retention rates at significantly lower costs. Applying speech strategically requires that a single solution – presenting a consistent voice and persona – be able to serve all caller needs, from greeting callers to giving them information, from conducting transactions to transferring calls and information smoothly to live agents. And the solution must do this in a way that is extremely cost efficient – meaning it costs significantly less than a live agent to deploy and operate, and it allows for quick and frequent changes.

Very recently, a new generation of speech applications has emerged, one that for the first time can deliver on this strategic vision. These new generation applications are delivering fully on the promise of improving both customer service levels and costs simultaneously. This paper outlines the technology breakthrough that allows new generation speech applications to be deployed strategically across the spectrum of call types, the potential of these new applications to dramatically affect business in a way traditional speech applications could not, and how to distinguish the new applications from the old.

Before You Buy: 2 Critical Questions You Need To Ask

Speech automation applied strategically has the potential to materially impact the business. It has the potential to deliver significantly higher customer acquisition and retention rates at significantly lower costs. So before you buy, ask 2 critical questions.

“Call centers around the world are using speech recognition solutions to automate customer interactions. Most of these are point solutions.”

Gartner

#1 What is the Speech Application Performance – Ability to Handle a Broad Spectrum of Call Types?

It needs to be able to serve your needs across a broad spectrum of call types while giving callers a consistent and satisfying experience. True application performance is measured by the spectrum of call types that can be handled with high customer adoption. A speech application cannot have a fundamental impact on the business if it just handles, say, store locator calls. It needs to be able to also service customers calling to request a catalog, check order status, place an order, check on loyalty program balances, and the like.

#2 What is the Speech Application Efficiency – Time to Deploy and Ease of Maintenance?

It must make business sense. The time and cost to implement must be less than what would be required to hire and train a new group of agents. It can't take three months to ready the solution to service each new call type. It can't take six months to reach payback for each call type. The cost to operate must be significantly lower than the expense of a live agent. And it must allow for immediate action on insights into customer behavior and on the constant flow of changes made to tune and improve customer service.

The Speech Application You Want to Buy

A speech solution such as this paints an appealing vision for the business. Imagine having a high-performance application that cost-effectively expands your call center capacity, precisely when you need it. Imagine new groups of virtual agents that can step in and handle any call type the instant wait times increase beyond a certain threshold. Imagine full control over both service levels and agent productivity levels. Imagine being able to introduce a new service that increases customer satisfaction and your brand image, at a fraction of what it

would cost to do so using live agents.

Imagine being able to offer an optional new channel for existing services, providing callers with choices and peace of mind. Imagine having the option to offload all information calls from live agents, keeping wait times down for revenue-generating calls while at the same time increasing agent morale. Imagine deploying a solution like this in a matter of weeks, and changing it in a matter of days, as needs change. Imagine having a high-performance application like this that is so economical that the more the speech application is used, the more your costs are reduced.

Let's take a look at the different approaches in speech automation, and why traditional approaches have not been able to deliver on this strategic vision.

The Challenge of Conversation

To understand any speech application, one must first appreciate the problem these applications were designed to tackle. Customer service conversations are not linear and robotic. There is a high degree of variability in even a simple conversation. Customers call with partial information (such as knowing the departure city but not the departure time or flight number); they answer a question with multiple pieces of information when only one was requested (such as answering with the departure date and time when only the date was requested); they correct a mistake they made on an earlier answer (such as mistakenly requesting tickets for two adults/one child when they really needed tickets for two children/one adult). An application cannot predict the path a given call will travel, or what information the caller will or will not have.

Every speech application works on top of a speech recognition engine – a high-quality,

Imagine new groups of virtual agents that can step in and handle any call type the instant wait times increase beyond a certain threshold.

Imagine having a high-performance application like this that is so economical that the more the speech application is used, the more your costs are reduced.

commoditized component that recognizes the individual words that callers speak. The speech applications extract meaning from the words and determine what action to take. Traditional speech applications can be segmented into two broad camps based on the approach they took to tackling the conversation challenge: the custom-development approach, and the toolkit approach.

Custom-Development Approach to the Conversation Challenge: Build All Call Paths

The custom-development approach provided usability but at a high cost. Speech vendors in this category approached the conversation problem by attempting to predict in advance all the paths a call could potentially take from each potential conversation juncture. They then coded the resulting explosion of call paths, for each customer service call type, so their application would be ready to service each call regardless of how the call played out. As a result, these custom applications were able to achieve a high level of user experience, but at a high price, and with an extremely long deployment time. In addition, because of the sheer volume of code that resulted, these applications were extremely inflexible – even a small change to the application was very time-consuming and expensive. Custom-developed speech applications had an extremely high capitalization cost, which meant businesses spent a large amount of money before the application ever serviced the first call.

An example of one of the more advanced custom-developed speech applications is Amtrak's "Julie." It took 18 months to get Julie to take her first call. At that point she told callers whether their trains were on time. It took an additional 12 months to get Julie to the point where she could answer fare and schedule questions. Finally, it took another 6 months to get Julie to accept credit card payments for tickets. It took a total of 3 years and \$4M dollars to get Julie to this level

of functionality, which then served Amtrak for a limited time only, as Julie was too difficult to change when business needs changed.

Amtrak's experience is representative of those who purchased speech applications that used the custom-development approach. The Gartner Group predicts that "60 percent of enterprises will be unable to capitalize on earlier investments in speech recognition."²

SLM: A Variant of the Custom-Development Approach

Statistical Language Modeling, or SLM, is a common technique used by many custom-developed solutions to "learn" the full range of possible answers to a question. In order to accomplish this learning it needs an input set of 10,000-100,000 sample utterances, which can take months to collect. As such SLM is too cumbersome and slow to use for every step in a call flow. Moreover, every time a question is changed, a new SLM sampling must be initiated. This is a time-consuming process which does not keep up with the speed of business change.

In summary, Forrester Research states that "it costs an average of \$300K-\$700K to deploy a custom application, and takes an average of 6-12 months."³

Toolkit Approach to the Conversation Challenge: Fixed Call Paths

The other traditional approach to the conversation challenge was the toolkit approach. This approach favored improving call center costs at the expense of service quality. Speech vendors in this category approached the conversation problem by tackling the most basic call types only, much like the touch-tone IVR vendors. Their toolkits delivered a structured development environment that reduced both the expertise and time required to develop code. The applications built by these toolkits were

The Gartner Group predicts that "60 percent of enterprises will be unable to capitalize on earlier investments in speech recognition."

Forrester Research states that "it costs an average of \$300K-\$700K to deploy a custom application, and takes an average of 6-12 months."

limited in functionality. Only simple tasks could be performed well, such as finding the store nearest to the caller's location. In the case of this simple task, the application would ask the caller for their zip code and then respond with the address of the nearest store.

Only a narrow set of call types are truly simple, however. The toolkit model broke down trying to accommodate call types that might seem simple – such as flight reconfirmation calls. With a flight reconfirmation call, the customer is calling to confirm that she is booked on a particular flight, or she is calling to get flight details such as flight number and arrival time, or both. If the conversation travels down the simplest and most desired path, it can be serviced by the toolkit-developed application. With this desired path, the caller would have her numeric confirmation number, the application would then look up the flight using the confirmation number as the record locator, and confirm the flight details back to the caller. However, as is often the case, the caller would not have the desired record locator. In this case, the application would need to take an alternate action. This alternate action could be to then ask for the flight number. The caller, however, may not know her flight number, or she may give the wrong flight number. Again, the application would need to know what to do. Even with a seemingly simple call type such as this, at some point a toolkit-developed application broke down – it would either start repeating a question and waiting for an answer it recognized, or it would transfer the call to a live agent. It would either force the caller down a linear line of questioning, or it would fail to bring the call to resolution.

Toolkit-developed applications handled a very limited scope of calls in a very limited way. The more sophisticated the interaction, the less likely it could be served with a toolkit-developed application.

Traditional Approaches Cannot Give Performance and Efficiency

Because of their inherent cost or performance limitations, traditional speech applications were used as point solutions. They were deployed either to perform a complex task at a high cost (where cost prohibited broad deployment), or to perform a simple task at a low cost (where performance prevented broad deployment). Without the ability to apply the solution broadly, the potential value of speech automation was never realized.

The New Approach: Navigation Skills to Traverse Call Paths Dynamically

Vendors with a radically new approach to tackling the conversation challenge have recently emerged on the scene and their applications are delivering fully on the promise of speech: they are dramatically improving both customer service levels and costs simultaneously. This new approach tackles the problem from a different vantage point. Instead of viewing the problem as the clarity with which all the possible call paths are predicted and the tools to then build these paths, this new approach calls for building the skills required to gracefully handle the conversation no matter what path it travels.

Instead of building call paths, then, this new approach involves building the skills needed to respond to any turn the conversation takes. As such, these new applications focus on the leverage point in the equation. Powered by these navigation skills, new generation applications can flexibly and gracefully traverse any call path. In this new paradigm, flexibility of navigation rather than custom-coded call paths allows the application to gracefully handle any call. And since the application comes with these navigation skills already built, there is not the time, expense, or baggage of custom code.

The application simply needs to be tuned to the processes that constitute your brand image, and configured to find your data.

Applications that use this new approach not only offer the highest performance yet, but they are extremely efficient, meaning they are quick to deploy and easy to change. This allows them to be applied strategically in the business, across the spectrum of call types.

With a myriad of speech applications on the market, how can you make sure to choose one that can be deployed strategically throughout your call center, and that will deliver the dramatic business impact you seek.

Measuring Performance & Efficiency Before You Buy

Performance and efficiency are the critical attributes that must both be present for a solution to be deployed broadly. These are the attributes that distinguish the new strategic applications from the old point solutions. The task, then, is to find a way to measure a solution's performance and efficiency before you buy. Let's take a careful look at how to measure these critical attributes.

Measuring Performance

True application performance is measured by the spectrum of call types that can be handled with high customer adoption.

Broad Spectrum of Call Types

For any speech application to deliver dramatic business impact, it needs to be able to handle all customer service call types – not just the simple information calls, but the more complex transaction-based calls as well. Regardless of how broadly you plan to deploy the speech application in the near-term, you need the option to deploy it across virtually all call types in order to ultimately

reap the benefits of strategic speech. If you lock yourself into a solution that doesn't perform across the spectrum of call types, you lock yourself into a solution that can't serve you strategically.

The first step in assessing performance is to make a list of all your call types. The following illustrates some of the calls that may appear on your list:



For any speech application to deliver dramatic business impact, it needs to be able to handle all customer service call types – not just the simple information calls, but the more complex transaction-based calls as well.

High Customer Adoption Across All Call Types

The second dimension of performance is customer adoption. An application is only high-performing if it is widely accepted by your customers. Customer adoption is your barometer of the customer's experience during the call, and how the customer feels when the call is finished. The best single measure of customer adoption is call completion rate, that is, the percent of callers that completed the intent of their call. This is the inverse of call abandonment rate. For example, what percent of those who called for flight status listened to flight arrival information? What percent of those who called for hotel reservations reserved a room?

An additional measure of customer adoption is that of conversion rate, meaning the percent of calls that were converted to a purchase. For example, what percent of those who called to explore products actually made a purchase?

Taking your list of call types, the next step in assessing performance is to calculate your benchmark customer adoption for each call type. What constitutes high customer adoption? For a solution to be deployed broadly and strategically, it must have performance that is comparable to that of a live agent.

For example, for calls you currently service with live agents:

- What percent of those who called to locate a store got store location details?
- What percent of those who called catalog request actually ordered a catalog?
- What percent of those who called to find a hotel actually booked a room?
- What percent of those who called to explore fares actually booked a flight?

For each call type, you need a call completion benchmark and where relevant, a call conversion benchmark. Your current performance is your best benchmark.

Measuring Efficiency

When a speech application is efficient, it is quick to deploy and easy to change. This gives you a solution that is both cost effective and nimble. Let's take a closer look at deployment and change.

Deploy in Weeks

Speech applications require a unique deployment effort for each call type they service. This is because each call type has a different purpose with different information gathering needs. The deployment effort will vary depending on the complexity of the particular call type and the complexity of the backend systems accessed (i.e. web-enabled vs. legacy, etc). To service a store locator call, an application only has to gather zip code or city/state from the customer, and then retrieve the nearest store locations from a single backend system. This is an example of the simplest call type and deployment should be very fast (6 weeks or less). The information required for an airline flight

booking, on the other hand, can be extensive and include a myriad of variables, some which are fixed and some which are flexible. Example variables include date, time, and airport of departure, and date, time, and airport of return. One customer may be flexible on selected variables in order to get a better price, while another customer may be flexible on different variables. In addition, the flight and fare data may be stored in multiple backend systems. This call type is extremely complex and will require the longest deployment effort (up to 12 weeks). Regardless of the complexity of the call type, deployments must be short in order for the application to be used broadly. If it takes four to six months to deploy an application for each new call type, the time and expense make it impractical to deploy the solution to service more than a select few call types. In order for the application to be deployed strategically, it should take about six weeks to deploy a simple call type, and no more than 12 weeks to deploy the most complex call type.

In addition, the initial deployment should establish a foundation for all future deployments. That is, there are business rules and data sources that are common across all call types. An example business rule would be a rule that states that the lowest priced hotel room is always listed first. An example of common data is the set of abbreviations that airlines use to denote airports around the world. You should only have to configure the application for common business rules and data sources once. So, once the application is deployed for the first call type, future deployments of like-complexity call types should take *less time than the first*. If, for example, the initial deployment is for the most complex call type, the initial deployment may take up to 12 weeks while all subsequent call type deployments should take less than 12 weeks.

Change in Days

Providing excellent customer service requires a constant flow of changes. The more

For a solution to be deployed broadly and strategically, it must have performance that is comparable to that of a live agent.

immediately these changes go into effect, the more responsive your business and the faster you get desired results. The reasons for changes are many: call centers change business rules to improve processes, they gain new insights into customer behavior, and they deliver promotions to motivate behavior. Some examples:

- A new business rule: a company's decision to ask customers to confirm that they understand the cancellation policy.
- A new customer insight: an airline that learned that the majority of callers were giving up upon receiving a third price quote. Instead of offering a 3rd fare, the airline began offering a fresh option – asking the caller if he or she would like to try a different route or departure airport.
- A new motivational program: a hotel that decided to offer a special discount for an immediate booking on a reservation.

To stay competitive, call centers need to be able to implement these changes in days, not weeks. An airline recently identified an improvement to its flight reservations call flow that had the potential to increase bookings. The change involved offering fare-conscious callers the option of checking prices on flights that departed a day earlier and a day later than the caller's optimal departure date. This fairly complex addition to the call flow was implemented in seven days. Even complex changes such as this are frequently needed, and you need the ability to implement them in days, not weeks or months.

Putting the Data to Work: Three Reference Phone Calls

Now you are equipped with the data you need to assess whether any speech application will have the performance and efficiency you need. The final step is to put this data to work. The only way to know that a given solution will perform to your standards is to talk to others who are using the solution in production. Ask the vendor for

three customers who are using the application across the broadest range of call types. Call and speak with these three references. Explore the following critical areas:

- Broad Spectrum of Call Types: Ask the references to list all call types the application is servicing. Moreover, ask them to list any call types the application is not servicing. Explore the difference between these two lists. Is the application deployed to service the full spectrum of calls? If not, why not?
- High Customer Adoption Across All Call Types: Explore the customer acceptance for each call type. Ask if the reference tracks call completions and conversion rates. How does the application perform on these measures for each call type? How does this compare to your live agent benchmark for each call type?
- Deploy in Weeks: Ask the references specifically how long it took to deploy the application for each call type.
- Change in Days: Think of changes you have made to your call flows and business rules over the past months. When calling references, discuss these changes with them. Ask if they have made similar changes, and specifically how long each change took to implement.

Review the answers from the vendor's three best customers. Do they paint a picture of the application deployed across a broad range of call types with high customer adoption? Are the results sufficient to ensure that the application will receive high customer acceptance across virtually all your calls? Will you be able to deploy every new call type in weeks, with deployment for the most complex call taking no longer than 12 weeks? Will you be able to make changes to call flows in days?

If the answers to the above questions are Yes, you are ready to take the final step in your evaluation process: listen for yourself. Call the reference call center, and complete transactions on their system as if you were one of their customers.

In order for the application to be deployed strategically, it should take about six weeks to deploy a simple call type, and no more than 12 weeks to deploy the most complex call type.

The only way to know that a given solution will perform to your standards is to talk to others who are using the solution in production. Ask the vendor for three customers who are using the application across the broadest range of call types.

Summary

Before you buy speech self-service, find out about a speech application's performance – ability to handle a broad range of calls. Find out about a speech application's efficiency – the time it takes to deploy and the ease of maintenance. New generation speech applications are delivering fully on the promise of improving both customer service levels and costs simultaneously. These applications change call center economics. The more the application is used, the more both service levels and costs improve.

Speech application performance is measured by the spectrum of call types that can be handled with high customer adoption. Speech application efficiency is measured by ease of deployment and change. The only way to know whether a given solution will deliver high levels of both performance and efficiency is to talk to others who are using the solution in production. For any application you consider, it is recommended that you ask the vendor to speak to their three best customers, and ask those references a list of questions related to the performance and efficiency they have experienced. If this critical evaluation step is taken with care, the solution you choose is likely to deliver dramatic business results for years to come.

¹ Gartner Research Note:

Take an Integrated Approach to Speech Technology, November 2003

² Gartner Research Note:

Take an Integrated Approach to Speech Technology, November 2003

³ Forrester:

Speech Applications Go Mainstream, March 2004