



Speech Rec: Temporary Slowdown, Faster Adoption Ahead

Advanced speech recognition, otherwise known as speech rec, is one of the most effective tools organizations are increasingly using to reduce customer interaction costs while improving quality service and with it customer retention and revenues. These solutions' without-wait, human-resembling dialogues have proven to keep more calls in automated voice systems, increase responsiveness to outbound notifications and shorten live agent calls compared with DTMF, a.k.a TouchTone IVR in most cases.



While speech rec deployment has slowed down with the recession there is a clear road ahead for more of them to be adopted. Yet hurdles remain before their applications become widespread.

Joe Outlaw, principal analyst, Frost and Sullivan reports that a majority of larger firms want to build automated customer contact applications with speech rather than with DTMF. They have recognized enough of the advantages, tested them out and now are waiting for the economy to strengthen more before greenlighting the projects.

Smaller businesses and contact centers have in contrast been much slower to deploy speech rec because they perceive the technology as still too complex and expensive, he points out. Application techniques and pre-programmed/configurable products have to get better before these organizations sign the dotted lines. There also needs to be more speech application developers who are willing to work on smaller projects.

"The fundamental business drivers for speech applications are still there," says Outlaw. "As the economy recovers they will drive decisions to resume speech application projects: at least for larger firms."

The speech rec technologies are improving, becoming less expensive and their implementation times are shrinking. Yet there is still often a gap between buyers' expectations and the realities of what the applications can deliver.

"We are time and again surprised when even companies that have deployed speech recognition ask us if the technology has improved to a point where they can simulate human-like conversational interaction with a speech system to completely replace the contact center agent," explains Fakhri Karray, CEO and co-founder, Vestec. "Firms are typically thinking of StarTrek and we have to point out that that kind of interaction is still far into the future."

Costs, Deployments and Budgets

Where the rubber meets the road is that potential speech rec buyers/clients have tight budgets and stringent timeframes for investments to show ROI. Yet the technology is not a quick or inexpensive fix. The applications can cost tens to hundreds of thousands of dollars and can take up to 18 months to deploy.

"Clients continue to drive requirements for an accelerated time to market for speech applications," reports Dave Pelland, director, design collaborative, relationship technology management, Convergys. "They're pressuring vendors to drive down the cost and time to market for application development and deployment to gain ROI quickly. They continue to view the cost of speech as too high for perceived value, requiring vendors to demonstrate strong ROIs. In many cases, projects are split up into multiple deployments to get some ROI benefits as quickly as possible. Leveraging performance guarantees and assurance of ROI are [also] key to deploying speech solutions."

Speech rec customers are picking easier, quicker-to-realize ROI speech applications. Michael Perry, product management director, contact center technologies, Avaya, reports that more firms are using speech to shorten calls, reach the right agents and in outbound notifications, all of which cost less and show positive results in less than 12 months. Another growing application for speech recognition he is seeing are as biometric identifiers on customers' accounts.

Speech-recognition vendors are also making their offerings more attractive through new partnerships. For example, Avaya now offers Loquendo's ASR speech engine in addition to IBM and Nuance products for its customizable Avaya Voice Platform.

"Loquendo offers a lower pricepoint for their core speech engine, which is available in many languages, and is popular



with companies, consultants and integrators internationally,” explains Perry. “Loquendo therefore gives our customers more choice in speech technologies.”

Digium, which created and owns the Asterisk open source telephony software, and Aumtech have partnered to provide users with Aumtech’s Media Resource Control Protocol Connector utility for the Microsoft Office Communications Server 2007 Speech Server. It features server-based licenses that support 48-plus ports and unlimited grammars and over a dozen languages.

“Our Asterisk customers have long recognized the value of speech recognition applications; however, most have delayed implementing these technologies due to the high entry costs,” says Bill Miller, Digium’s vice president of product management. “Aumtech’s solution is the lowest-cost at higher capacities for world-class speech technology and may prove to be the ‘tipping point,’ finally bringing speech to the masses.”

One option companies are considering to lower costs is buying packaged applications such as Nuance’s SpeechPak Application Kits, reports Frost and Sullivan’s Outlaw. These tools can save 10 percent to 15 percent or more off pure custom installations.

There is increased customer demand for hosted applications. Outlaw says with hosting organizations do not have to outlay capital for, manage or run the applications. The agreements also create built-in incentives for success as the hosts are often paid by completed-call-minutes.

“Consequently the hosts have the incentive to tune the applications – for improved operation and increased usage,” says Outlaw.

Jeff Foley, senior manager of marketing for Nuance’s enterprise division is seeing more of his firms’ customers shift to hosted environments from on-premise to realize speech rec benefits with minimal involvement with hardware and software upgrades, system designing and fine-tuning.

“A lot of companies don’t want to deal with those details and are willing to pay someone like Nuance to take ownership of the system and provide continuous improvement,” says Foley.

Hosting is making speech rec more practical for applications such as outbound notifications. Speech offers value-add, explains Grant Shirk, director of industry solutions at Tellme Business Solutions at Microsoft Corp., because it enables more compelling, efficient campaigns, for early-stage collections, alerts, or customer care scenarios.

“Because the scope of outbound applications is much smaller than inbound –they are usually focused on completing a very specific task – on-premise speech deployments were often cost-prohibitive,” explains Shirk. “However, with the advent of on-demand speech as a service, more companies now have access to the technology in an affordable way.”

Refining but not Revolutionizing Technology

Speech rec technologies and deployments continue to be refined rather experience breakthrough developments. The net results are steadily increasing completion rates, shrinking costs and shortened install times.

Avaya is adding new functionality to Dialog Designer that enables developers to quickly build contact center workflow and routing strategies using the same tools they use to build VoiceXML self service applications. Developers working on Avaya contact centers only need to know one tooling environment, explains Perry. They can create rich user experiences by tightly integrating self service and agent assisted transactions quickly and easily.

Avaya’s self-service platform will be leveraging several core technologies from the Nortel acquisition. The enhanced solution will be incorporated into Avaya’s new contact center platform, Next Gen Context Center, expected to be gradually rolled out this year.

“The new portfolio allows us to incorporate some of the advanced media processing and tooling capabilities from Nortel that will allow us to better integrate with a wider variety of contact center environments,” says Perry.

Frost and Sullivan’s Outlaw sees greater application development tool availability and improvement; as they become better the more speech applications tend to get built. There are more industry standards such as Voice XML which enhances and makes it easier and more desirable for companies to build speech applications. Analytics tools are also being used to discover opportunities to improve existing speech apps and build business cases for new ones.

He also sees steady recognition rates and language domains improvements. He points to VoltDelta’s CrystalWave, which takes outputs from multiple recognizers running against the same conversations, applies context analysis and uses that to boost recognition rates over what they would get out of any individual recognizer.

A recent white paper by VoltDelta adds that the recognition results can be immediately refined through a speech process known as robust parsing. This technique works to verify results that appear consistent with the data set or discounting those that fall out of logical parameters.

Outlaw also points to Nu Echo’s NuGram Platform, which helps developers author, tune and manage speech grammars. Nu Echo’s NuGram has also been integrated with Voxeo’s VoiceObjects Service Creation Environment. This allows developers to easily create either static or dynamic grammars that can be used with VoiceObjects technology to efficiently build multichannel self-service applications.



Vestec's Karry reports that artificial intelligence-based algorithms are significantly improving recognition quality for both native and non-native speakers, which in turn is increasing customer satisfaction with speech applications. Advanced noise-cancellation techniques are also helping improve recognition quality in noisy mobile and VoIP environments. At the same time, there has been a dramatic decrease in prices of speech recognition software. High quality, standards-based speech recognition engines can now be licensed at less than \$100 per channel, "a figure that makes speech recognition truly affordable for the first time to the vast majority of smaller business and enterprise markets," says Karry.

In the next year, Tellme customers can expect to receive enhanced capabilities for speech-enabled outbound, new performance optimization tools to help them improve task completion rates, continued core engine improvements, and the launch of cloud-based routing and queuing services to optimize contact center resource utilization. It plans to debut technology that will open up its speech platform to more channels, enabling speech interfaces for mobile, online and other devices.

There continues to be a slow shift to natural language as its processing improves from directed dialogue, which currently dominates speech rec applications. Natural language-based applications are becoming more conversational and therefore able to keep users in the automated systems longer, making them to become more desirable for companies to implement. Yet they are still more expensive and time-consuming to deploy than directed dialogue and are still a ways yet from making talking to machines as "natural" as conversing with people.

"Having a computer with enough built-in alternatives to what might come out of someone's mouth if you 'how can I help you today' it is going to take a lot longer before we have enough grammar and vocabulary to come back with a rational answer," says Outlaw.

Making speech applications in directed dialog and natural language more accurate, reliable and feasible is the rise of distributed computing resources such as Microsoft's Azure, reports Tellme's Shirk. Harnessing this computing power, he says, "promises more accurate recognition, greater automation per task, and higher user satisfaction."

Nuance's Foley thinks the ideal applications are those that give customers choices via directed dialogue, which is more straightforward and less complex than natural language but have natural language capabilities underneath. This way if a customer goes out of grammar the application can understand what they say and bring them back to the conversation.

Nuance is improving its applications to better understand the human response, i.e. individuals saying something that the systems are not designed to expect. It is exploring ways to

design the system and taking advantage of enabling technology advancements such with filters, contextual analysis, fuzzy matching and parsing information.

"The challenge is how can make our technology better at listening to what end-customers are saying and matching them to the possible options they have," says Foley. "In many cases we can understand exactly what customers are saying, we can transcribe it to go back to what they've said, but how do you map that to a choice on the speech rec menu?"

Customer Acceptance

Equally if not more importantly than the application development work in enabling speech rec is greater consumer and business familiarity and acceptance. One of the literally big drivers is the growth of mobile communications including as alternative to having residential landlines. Speech solutions eliminate handheld operations, which are becoming prohibited behind the wheels of vehicles in a growing list of jurisdictions. A study by Forrester for Nuance found that a growing segment of mobile consumers prefer self-service wherever possible especially speech service.

"If you look at the way cellphone, smartphone and even landline phones both corded and cordless are designed today it is very inconvenient to shift back and forth between the microphones and receivers and the keypads," Foley points out. "Speech rec bridges that for automated applications."

Enabling customer acceptance are more consumer-oriented speech applications. They are increasing awareness how speech systems work and how they can be used to interact with companies. Foley cites Ford's Nuance-developed SYNC and Nuance's Dragon iPhone app as examples.

"Consumer familiarity with speech is a big win on contact center side, because it makes consumers more willing to use it if presented with it when calling," says Foley. "And consumers are beginning to appreciate companies who put good speech systems in place because it gives them a more positive experience as compared to waiting on hold for five minutes and then talking to an agent who can't help them beyond the script they have." **CIS**

The following companies participated in the preparation of this article:

Avaya
www.avaya.com

Aumtech
www.aumtech.com

Convergys
www.convergys.com

Nuance
www.nuance.com

Tellme
www.tellme.com

Vestec
www.vestec.com

VoltDelta
www.voltdelta.com

Voxeo
www.voxeo.com

Original Article: <http://www.tmcnet.com/call-center/0310/table-of-contents.htm>