

Right-Sizing Speech Analytics: Why Contact Centers Need a Scalpel, Not a Chainsaw



A Frost & Sullivan White Paper

EXECUTIVE SUMMARY

Speech analytics is a rich technology for improving customer contact operations and customer experience. It has applications in boosting agent performance and providing valuable insights into customer behavior. But adoption has been slow, because the first generation of tools overstretched – providing too many unneeded features that pushed price and complexity past the point where typical centers could afford to deploy it.

What contact centers need are speech analytics applications that are streamlined to use less processing power, and provide a more effectively tailored mix of features that have proven use cases and demonstrable applicability inside the center.

SPEECH ANALYTICS DEFINED

Speech analytics is best defined as a software and hardware application used in contact centers to do the following:

1. Aggregate pools of recorded voice calls from telephony platforms, especially from call recording systems.
2. Harvest significant trends and patterns from the data. It should allow a user to search for a particular set of patterns, as well as find patterns that the user was not explicitly looking for.
3. Compare the data across time and across data sets, presenting information at a micro level (i.e., what happened in this call) and at a macro level (i.e., what is happening in the call-base as a whole).
4. Relate the data to actionable steps in multiple parts of the organization to correct business process flaws.
5. And provide a mechanism for sharing the findings with relevant professionals inside and outside the contact center.

Clearly, "speech" only enters the picture at the beginning of the process, when the unstructured voice calls are identified and gathered as the source of the unknowns, and it remains in the background throughout the process, as a speech analytics tool should be best at finding connections between what is said in a call and the outcomes that stem from it.

Most observers have been amazed at the technological sophistication and power of speech analytics systems. What has happened to speech tools under the hood has been nothing short of a revolution: unparalleled recognition, computational understanding of timbre, emotion and tone. No one who has seen an automated system parse and analyze an interaction as complex as a voice call can be anything less than impressed. In short, speech analytics is the pinnacle of today's tech.

Even as we marvel at its capabilities, though, few companies have chosen to deploy speech analytics in their contact centers. It's a conundrum: if it works so well, why aren't contact centers using it? Where does the disconnect lie between the value of the technology to solve real world problems and the businesses it is designed to help?

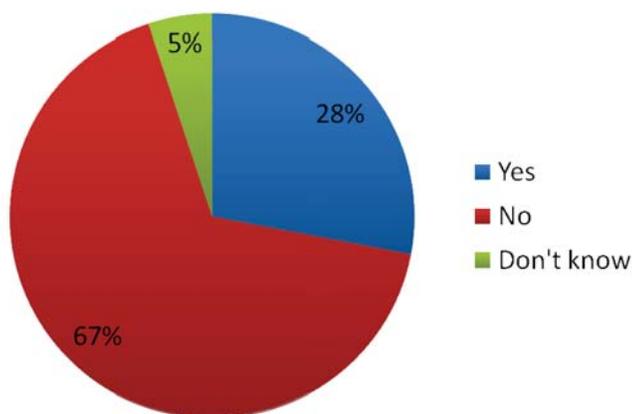
The disconnect arises when contact center professionals assess the ultimate value of speech analytics. In a recent survey of end users by Frost & Sullivan, center managers who were not using speech analytics were asked what had deterred them from the purchase. Among the most cited responses were a lack of clear business value for the tool (38%) and ROI that was too far in the future (23%). And the top deterrent to deployment was, unsurprisingly, cost (44%).

These responses say much more about the state of mind of contact center managers than they do about the actual value and capability of speech analytics software. They indicate a wide gap between the way speech analytics has been developed and brought to market in the customer care industry, and the actual day to day needs of that market. Speech analytics is, in fact, a transformative technology that can improve agent performance, boost the customer experience, impact revenue, and spread important information among many key decision makers within an organization. Those that can see its clear business value will emerge as leaders with a competitive advantage in their customer care infrastructure.

Among the most cited responses were a lack of clear business value for the tool (38%) and ROI that was too far in the future (23%). And the top deterrent to deployment was, unsurprisingly, cost (44%).

Use of Speech Analytics Systems

Q: Do you use SA systems in your contact center?
(2009, N=332)



Source: Frost & Sullivan End-User Survey, 2009

UNDERSTANDING THE DISCONNECT

Some of these perceptions are rooted in the way speech analytics has been presented to potential users over the last five to seven years. In hindsight, it appears that "speech analytics" was rolled out to the industry too early, with contradictory messages that touted the advanced technology of the most expensive and complex features without emphasizing the corresponding benefits. Emotion detection, for example, has long been considered a selling point of speech analytics and an emblem of how sophisticated the technology can be. But emotion detection in and of itself does not bring much value to a contact center – it is a wondrous demonstration capability, but without context and meaning and workflows underpinning the knowledge of a speaker's emotion, where is the hard dollar value to offset the high cost of deployment?

Before there was "speech analytics" there was just-plain analytics, a category of tools that transcended the contact center and impacted many parts of an organization. Financial departments were used to using analysis tools in the form of business intelligence and ERP systems. Marketers used analytics to glean data on some customer behaviors through a wide variety of tools and data sets. In contact centers, "analytics" circa 2003 was taken to mean the harvesting of data from the call routing switch and the quality systems and using them to understand (and hopefully improve) agent performance. Contact center analytics gradually evolved into today's Performance Management tools. In that role, they are a bit more focused on implementing improvement solutions than in ferreting out root causes of problems.

Also before there was "speech analytics" there was speech recognition. Automated systems to understand and act upon speech have been in use in contact centers since the late 1990s, and have now come to be a vital part of IVR and voice portal self-service strategies. Speech recognition was slow to achieve its success, though, because of the high cost of implementation, a slow-to-improve accuracy rate, and the need to continually maintain and update recognition libraries.

Eventually, speech recognition matured and is now the underlying engine powering today's speech analytics systems: you can't analyze what you can't recognize. But the confusion and overlapping meanings that persist around the term "speech analytics" have contributed to preconceived notions left over from both speech and analytics of the past.

It has become apparent that the speech component and the analytics component are truly separate elements. You can have an analytics system that does not incorporate speech processing; likewise, you can have a speech recognition engine that does little or nothing to analyze the content of the data.

Parallel to speech development was the widespread adoption of call recording and logging technologies in contact centers. What began as a compliance regimen in specialized centers quickly achieved broader acceptance as part of an operational formula for improving agent performance and customer interaction quality. When recording and storage became a commodity product, vendors of those recording tools tried to move up the value chain by offering performance-related analytical tools.

Analytics and speech were moving in synchronous directions. Analytics moved towards integrating the mountain of contact center data that was newly available. And speech moved towards having the sophistication to identify vocal patterns in the recorded calls.

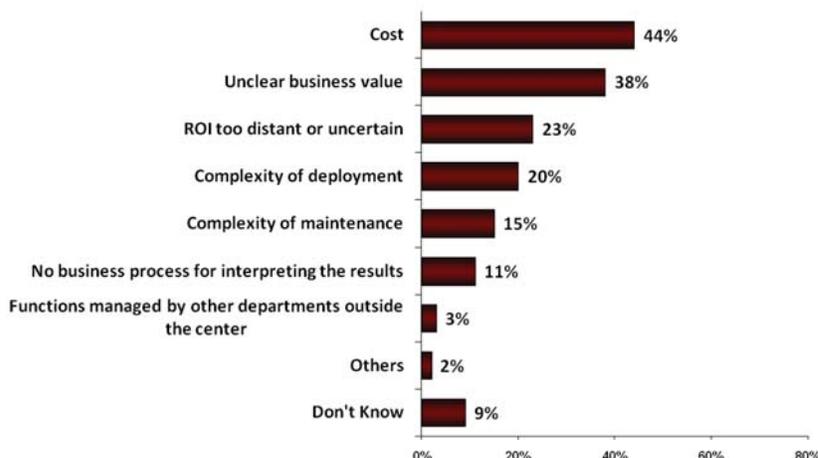
It has become apparent that the speech component and the analytics component are truly separate elements. You can have an analytics system that does not incorporate speech processing; likewise, you can have a speech recognition engine that does little or nothing to analyze the content of the data.

To take advantage of this convergence, vendors of quality monitoring software began to merge the two strands to mine the speech data and do “word spotting” and phoneme analysis.

When speech analytics was positioned to the contact center as a way of mining agent recordings for adherence to good quality practices, it failed to catch on because traditional QM tools are cheaper and easier to use than speech. And when speech analytics was positioned as a way to gain insight into customer behavior, the contact center balked because of the complexity and required capital investment in technology.

Barriers to Speech Analytics Adoption

Q: What has deterred you from purchasing this technology (2009, N=222)



Source: Frost & Sullivan End-User Survey, 2009

PROCESSING EATS RESOURCES

Misconceptions about the cost of speech analytics also come from a lack of understanding of the true hardware power required. Speech processing is very expensive to scale – the more calls you want to process, the more hardware you need to install. Unlike other workforce optimization tools, speech analytics has not been just a software purchase, which sometimes comes as a surprise to those buying it in tandem with a quality or workforce management application.

Software licensing is just one factor in the overall total cost of ownership, next to hardware, services and the IT costs of managing those resources. Many centers that consider a full blown speech analytics project turn away when they learn the full resource exposure needed when you commit to high-intensity features like emotion detection.

There are other features that drive up the cost through processing power requirements: full and/or separated text transcriptions of calls, for example, and needing every word separated between the agent and the caller. Those features are not needed to get the value out of analyzing voice calls.

Speech processing is very expensive to scale – the more calls you want to process, the more hardware you need to install. Unlike other workforce optimization tools, speech analytics has not been just a software purchase, which sometimes comes as a surprise to those buying it in tandem with a quality or workforce management application.

Managers often make the mistake of assuming that they need to analyze all their calls. Even if a center is recording 100% of its calls, it doesn't need to run them all through speech processing in order to measure quality, agent performance or customer experience. Critical trends can be spotted with just a sample of calls. The key value is in knowing how to select the best sample to get the right results. That pre-filtering is a software application with low processing overhead.

When you look at speech analytics in this context, you see a technology that has been feature-rich but solution-poor: a wonderfully advanced toolset with bells and whistles that have yet to demonstrate a consistent ROI case, even if one exists. This is the fault of many of the vendors, who frankly have not yet articulated the clear value proposition inherent in speech analytics. Until now, they have put complex systems forward into the marketplace without expressing where contact centers will find the clear, immediate benefits.

THE RIGHT-SIZED TOOL FOR THE JOB

The contact center industry has seen blue-sky technologies before, and been skeptical. Historically, the pattern laid out by CTI, CRM and IP telephony have all sketched a path that began with ambitious breakthrough technologies and bright promises of benefits. And in their first practical iterations they were somewhat disappointing. Early adopters often found they had overspent for features that had great appeal in the abstract, but little application in reality.

Critically, they also found that using leading edge technology can extract a stiff price in time and resources. Making tools like CTI and CRM work within the existing structures for handling customer interactions required extensive configurations, integrations, and consulting ventures that were often unexpected. None of those breakthrough technologies of the last 15 years had real impact on business processes until feature sets and deployment methods had standardized based on the specific practices needed by contact centers, pulled by demand rather than pushed by the suppliers.

The speech-centric approach to analytics that has prevailed for the first cycle of speech analytics development this decade has come mainly from companies that have originated outside the contact center industry. This has led to:

- an overly complex infrastructure of features that emphasize the power of the audio recognition technologies (emotion sensing and differentiating between speakers, for example)
- less of a focus on analyzing contact center activity emphasizing specific actionable steps that managers can use to find – *and solve* – problems.

Inside contact centers, speech is just one of many streams of important real-time and historical data that needs to be analyzed and applied to performance improvement. Speech is the one that is least-well understood, but it does not stand alone.

Unfortunately, the blue-sky approach to creating speech analytics applications has artificially skyrocketed the cost to end users, while providing little in the way of demonstrable improvement.

None of those breakthrough technologies of the last 15 years had real impact on business processes until feature sets and deployment methods had standardized based on the specific practices needed by contact centers, pulled by demand rather than pushed by the suppliers.

Contact centers are emphatically searching for tools that will help them improve agent efficiency and the customer experience. Frost & Sullivan research shows that Performance Management and Workforce Management software are being used by 67% and 58% of centers, respectively.

When asked what are the most important criteria they use in judging the effectiveness of their agent optimization tools, the most widely cited response was customer satisfaction improvement, followed by cost reduction and customer retention improvement. This survey showed that contact center professionals are responding to the downturn in the economy by seeking out solutions that have clear, explainable value.

And they have spoken with their wallets, that they are still willing to purchase tools like performance management, web analytics and quality monitoring, even when technology budgets are tight. What those tools have in common is that they don't require a new labor infrastructure to deploy or manage their solutions. They fit into the existing operational modes of standard contact centers.

What contact centers want, and need, is a speech analytics solution that

- a) is optimized for contact center needs in terms of the feature set and pricing – features geared to contact center improvement and not to blue-skying overall business issues
- b) has a footprint that can be managed by existing contact center personnel and doesn't require dedicated resources
- c) does not imply a commitment to an endless integration, customization and professional services cycle.

In other words, contact centers need speech analytics that is tailored to be the right size for their environments.

WHEN LESS IS MORE

So what would a right-sized version of such a complex application look like? An appropriate analogy can be found by looking back at the way workforce management software made headway in contact centers in the late 1990s and early 2000s.

In the mid 1990s, WFM was available from several vendors for contact centers at the very highest end of the market. The table stakes for deploying WFM in those days was upwards of \$100,000 (a sizable investment at the time), and it required its own dedicated staff to create schedules and maintain the system. The market was distinctly segregated between the contact centers with high volumes using the available WFM tools, and all the others who used very rudimentary shift and volume calculators that were little more than spreadsheets augmented with Erlang C functions.

The vast majority of centers did not and could not invest in the full-blown WFM systems until several independent vendors took the essential features of those high-end packages, stripped out the bells and whistles, developed graphical user-friendly interfaces for the front end, and created a mid-market product segment.

With more companies using more sophisticated tools to schedule their workforces, the craft of allocating labor resources became more professionalized. It is now seen as a necessary component of good management. Few argue that there is a "lack of clear business value" to WFM; the value of

What contact centers want, and need, is a speech analytics solution that is optimized for contact center needs in terms of the feature set and pricing – features geared to contact center improvement and not to blue-skying overall business issues

matching resources to calls is obvious. And there are countless case studies to verify the ROI value of using the tool to cut costs.

The key to right-sizing lies in excluding features that appeal to engineers but that don't have immediate practical value to an end user. In WFM's case that meant taking out the high-end ability to schedule across multiple sites and skills in order to introduce more users to the tool's core benefits: matching calls to headcount. That set a baseline for what features were "necessary" versus those that were optional.

A similar model, applied to speech analytics, would say that you could safely leave out some of the high-end features that are tangential to contact center operations, and by doing so, paradoxically add value to the application.

Some features are valuable mainly to long-time users who have already learned how to analyze the content of recordings for performance improvement. For example, emotion detection makes more of an impact when you are using analytics to interpret customer behaviors than improving agent performance. In most centers, emotion detection is not something that needs to be automated – it is enough to have an agent tick a box on a screen field to indicate for the record that a caller was "irritated" or "delighted," and much less expensive.

Similar dynamics apply to features like text transcription and real-time keyword alerts – they are value-adds that enhance the power of a speech analytics system. But they are not core to its function, nor core to its value proposition. Contact centers more realistically need speech analytics applications that are extremely lean, tools that are scalpels for precision adjustment of call handling and customer experience management. A right-sized application leaves room for feature growth up the value chain as time goes on. But first, it proves its value in a cost-sensitive, real-world environment.

PROVING THE CONCEPT: ENVISION INTERACTIONIQ™

The demand for a more user-friendly speech analytics system with a smaller footprint has not fallen on deaf ears. Envision Telephony, a Seattle-based company with long experience serving contact centers with powerful, yet easy-to-use agent performance optimization solutions, has crafted a solution aimed specifically at helping companies integrate speech analysis into the broader performance improvement portfolio.

In 2008, Envision's full-powered speech analytics offering (SpeechMiner) received the Frost & Sullivan Product Line Strategy Award for the company's efforts to deeply integrate the analysis of voice recordings into a coherent contact center improvement program. More recently, Envision has unveiled a new analytics tool, Envision InteractionIQ, with a narrower feature set and a simpler deployment model.

Envision InteractionIQ is an integrated part of the Envision Centricity™ agent optimization platform that includes the core elements of call recording, workforce management, quality monitoring and performance improvement, now standard industrywide. The analytics is not a standalone application, it's fully built in so that a contact center user can access the speech analytics directly from the same user interface she would see while managing quality or other aspects of agent performance.

A right-sized application leaves room for feature growth up the value chain as time goes on. But first, it proves its value in a cost-sensitive, real-world environment.

The most important element in right-sizing the application is bringing down the enormous overhead in processing power that larger-scale systems require. Traditional speech analytics works by offloading the entire recorded contents of a contact center into a dedicated processing engine. Instead, Envision InteractionIQ starts from the proposition that you don't need to analyze all that content. Not all content is relevant in any given context, so the first job of the analytics system is to help the user provide that context, through a smart filtering front end.

These days, a high portion of contact centers are recording 100% of their interactions because the cost of storage is so low. But the cost of processing all that data is higher than the cost of storing it. It can take two to three times as long to process audio as the length of the recordings themselves, depending on the accuracy desired. That's why it's so important to right-size not just the application, but the expectation of how much audio a user is actually going to need to process.

In most cases you don't have to pick up 100% of the utterances of particular words across a contact center's reservoir of calls. You need just enough to see where trends are occurring. Selective sampling works from both a training and a business perspective. Envision InteractionIQ lets users filter before the processor gets to work on the audio. The filtering tool is consistent with what users already encounter in their QM recording searches. It just adds a "content" element to the attributes that are searchable.

The entire schema of the filtering mechanism in Envision InteractionIQ is built to reduce and focus the processing load on the sample of calls needed to address specific business applications and solve particular problems. Users can filter based on agent, call type, customer value, date range, disposition code, or recording time, among other elements. Reducing the processing load reduces the application's overall footprint, keeping the cost low and the deployment simple.

Envision InteractionIQ incorporates a metadata feature called Smart Tags, which allows users to identify specific words that are important to their business, and the system will automatically mark the calls it processes that include those words. For example, in the wireless industry, the tags might include the names of competitive companies. And because it's only applied to the filtered call set, adding Smart Tags doesn't add to the processing load.

Envision InteractionIQ presents to the user the same way the rest of the Envision Centricity application suite does. So for a person working inside the contact center, trying to identify gaps in agent or center performance, using speech analytics becomes a natural extension of the way they would look at other forms of data to isolate issues and find solutions.

For example, in Envision Centricity, users can "subscribe" to calls based on criteria like call duration or call disposition, receiving them routinely within their inbox. With Envision InteractionIQ, speech content becomes just one more analytic criteria, just as easy to filter into the mix and navigate through. It builds a context for the use of speech. And that context extends outward to enterprise users, who now can automatically receive call samples where customers mention specific words, for example, both "cancel" and a competitor's name.

In fact, speech analytics is at its most valuable when, as in Envision InteractionIQ, a user can combine any piece of data in the contact center universe, whether it comes from an ACD, CRM system or QM data. You can create some pretty sophisticated models that deliver reports and/or the calls themselves

In fact, speech analytics is at its most valuable when, as in Envision InteractionIQ, a user can combine any piece of data in the contact center universe, whether it comes from an ACD, CRM system or QM data.

because in Envision InteractionIQ all that data is already integrated and pulled from the audio vault based on the Smart Tag capability.

What is revelatory about Envision InteractionIQ is that it proves speech analytics can be streamlined without being dumbed-down. It offers functionality for as sophisticated an analysis tool as a contact center needs. It reduces the processing load to keep down the initial (and total) cost of buying speech analytics. Essentially it is addition by subtraction, because without that subtraction few centers have shown a willingness to add speech to their repertoire of data analysis. Envision's strategy has been to build specific new features and modules into the suite in response to customer needs. At several points in recent years the company has rolled out components of the software suite that address very specific pain points in the contact center that would only be apparent by carefully observing and collaborating with customers.

And this is first and foremost a contact center improvement tool – it slides neatly into the existing optimization suite that tightly integrates all the contact center data sets and operational functions into one interface. Because Envision InteractionIQ is tightly integrated with Envision's Click2Coach (quality monitoring and e-learning) solution in a simple Web interface, it's easy to formulate and quickly take coaching action when issues or trends are identified. This makes Envision InteractionIQ an especially actionable and proactive analytics tool, versus just another performance metrics chart on a dashboard. It is a highly advantageous way to get started exploring this important (and neglected) data set without overpaying for advanced features and processing power that is of little immediate use.

THE TAKE-AWAY

Contact centers looking for insights into their customer interactions should strongly consider speech analytics as a tool for seeing inside their calls. Speech analytics can point the way to root causes of problems and suggest solutions. It is a valuable way to boost agent performance by finding trends. In considering speech analytics, contact center managers should keep these ideas in mind:

- It is not necessary to subject 100% of calls to speech processing, even if you are recording 100% of calls. It is more important to select a well-targeted subset for processing.
- Processing requirements are the primary driver in raising (or lowering) the cost of speech analytics deployments. Hardware should be front and center in feature selection.
- Speech analytics is no longer an "all or nothing" proposition: vendors are responding to the call for streamlined, right-sized applications that do the job without foisting unnecessary overhead onto a center's capital expense budget.

What is revelatory about Envision InteractionIQ is that it proves speech analytics can be streamlined without being dumbed-down. It offers functionality for as sophisticated an analysis tool as a contact center needs.

CONTACT US

Beijing
Bengaluru
Bogotá
Buenos Aires
Cape Town
Chennai
Delhi
Dubai
Frankfurt
Kolkata
Kuala Lumpur
London
Melbourne
Mexico City
Milan
Mumbai
New York
Oxford
Paris
San Antonio
São Paulo
Seoul
Shanghai
Silicon Valley
Singapore
Sydney
Tel Aviv
Tokyo
Toronto
Warsaw

Silicon Valley
331 E. Evelyn Ave. Suite 100
Mountain View, CA 94041
Tel 650.475.4500
Fax 650.475.1570

San Antonio
7550 West Interstate 10, Suite 400,
San Antonio, Texas 78229-5616
Tel 210.348.1000
Fax 210.348.1003

London
4, Grosvenor Gardens,
London SW1W 0DH, UK
Tel 44(0)20 7730 3438
Fax 44(0)20 7730 3343

877.GoFrost
myfrost@frost.com
<http://www.frost.com>

ABOUT FROST & SULLIVAN

Frost & Sullivan, the Growth Partnership Company, partners with clients to accelerate their growth. The company's TEAM Research, Growth Consulting, and Growth Team Membership™ empower clients to create a growth-focused culture that generates, evaluates, and implements effective growth strategies. Frost & Sullivan employs over 45 years of experience in partnering with Global 1000 companies, emerging businesses, and the investment community from more than 30 offices on six continents. For more information about Frost & Sullivan's Growth Partnership Services, visit <http://www.frost.com>.