




Blending Speech Analytics with Traditional Quality Management

White Paper



*Speech Analytics
provides contact
centers with a
technology that can
process 100% of
calls and search for
sentiment and specific
speech patterns.*

OVERVIEW

Traditional Quality Management is based on manual evaluations. It relies on evaluators, usually supervisors, to listen to calls and score them according to fixed criteria. It provides accurate, measurable analysis of agent performance over time, and can be used to direct coaching efforts.

Speech Analytics, which is rapidly becoming a mainstream technology, provides contact centers with a technology that can process 100% of calls and search for sentiment and specific speech patterns.

Is one better than the other? Let's compare:

Criteria	Traditional	Speech Analytics
Evaluation Method	Human	Computer
% of calls	< 5%	100%
Speed	1x - 2x Call Length	<< Call Length
Accuracy	High	Varies

At first glance, we can see that Speech Analytics has some key advantages: it can analyze 100% of recordings without human labor. This is an extremely powerful proposition for virtually every contact center. Its weakness is that accuracy, or quality, is a variable that needs to be closely examined before the technology can be effectively deployed. A Speech Analytics consultant we know likes to say, "Don't confuse Speech Analytics data with hard data." What does that mean? Let's look a little closer...

“GROOMING” A SPEECH ANALYTICS SOLUTION

Speech Analytics solutions use Artificial Intelligence techniques to analyze a recording. They don't “listen” to a recording like humans do – they actually “parse” a recording, after breaking it down into tiny bits of speech (hence the term Speech Analytics).



Computers can't interpret or decipher text the way the human brain can. But what it can do – very quickly – is compare a small piece of unidentified speech to a bunch of other, equally small but identified pieces of speech.

The first key to Speech Analytics is to understand this parsing action. A speech segment – a word or phrase – is broken down into what's called phonemes; which are defined as “perceptually distinct units of sound.” Artificial Intelligence is then used to match phonemes to the samples it's been taught, in its model or “dictionary.”

The way that Speech Analytics systems are “trained” to recognize speech is through recording samples. The more samples provided, the better the system will recognize words and phrases on its own. There's the rub – the Speech Analytics system is only as good as the samples it's been “taught.” To improve accuracy, the system needs a wide variety of samples from which to compare. It needs to be able to separate the speech signal from the noise, and it needs to be able to match the signal to a reference sample. Therefore, the two main challenges that must be overcome for the Speech Analytics system to produce a reliable speech-to-text transcript are:

- The system must have an adequate sample library in which to compare
 - The system must be relatively free of noise
-

Artificial Intelligence is just that – artificial. Computers can't interpret or decipher text the way the human brain can. But what it can do – very quickly – is compare a small piece of unidentified speech to a bunch of other, equally small but identified pieces of speech, and quantify how well it matches the known samples. Then it can use that analysis to pick the best match and label the unidentified piece. It can do this on many levels – phonemes, words, phrases, and sentences – and it can make multiple passes to further improve accuracy. Speech Analytics systems can also track the probability that the match is accurate – another thing humans can't do very well.

Asking “How accurate is your Speech Analytics engine?” is like asking a mountain climber how fast they climb. The answer is always another question: “Which mountain?”



The samples used to “train” the system, therefore, are a critical part of the total solution accuracy. A model that uses samples in English will have very low accuracy processing a German speaker. Language, accents, dialects, and colloquialisms present challenges that can be overcome if and only if the training samples have been carefully selected to provide enough coverage to allow the system to make a match with an acceptable probability. Since we're talking about contact centers, we also need to account for vocabulary. Contact centers will have their own unique vocabularies, based on the company, its business model, and the industry it serves. A hospital will have a very different vocabulary than a cable TV operator. Brands, product names, and other terms unique to a specific contact center will not have been encountered in the typical Speech Analytics model.

THE CHALLENGE WITH SPEECH ANALYTICS

The challenge of providing applicable training samples can be overcome using various methods. Custom models can be developed using previously recorded interactions as training samples. Many novices to using Speech Analytics don't realize that an out-of-the-box solution isn't tuned to their enterprise – and that much work (and expense) will be required to do so.



The second major challenge is to make sure that the system is free of noise, so there's no interference to the "signal" (phoneme, word, or phrase) that the system is trying to match. Having a noise free telephone recording environment is like asking an infant to eat baby food without making a mess: it's not going to happen. Telephones are themselves noisy – narrow-band audio requires high compression, handset and headset microphones are made for robustness and cost more than fidelity, and chances are extremely high that at least one-party in a two-party or multi-party call is using a cell phone. And that doesn't include issues like background noise, stammering, coughing/sneezing, guttural sounds, speakers talking at the same time, or other noise sources.

Speech Analytics has to be told what is meaningful or not meaningful for a given contact center.

Accuracy is a function of the engine, to be sure, but more to the point, it's a function of how well the engine can match a particular audio sample to its library of "learned" samples, given the noise in the sample. You can't assign an accuracy value to an engine anymore than you could assign a speed value to a mountain climber. The answer is always another question – "which mountain?"

SECOND-LEVEL GROOMING

Speech Analytics systems must be groomed on another level as well. Remember that Speech Analytics is a programmed search; the system is looking for meaningful words and phrases. The key word in that sentence is **meaningful**.

In traditional QA, a supervisor doesn't require a list of search words or phrases to analyze a call. Their own training and experience is sufficient; the evaluation form (and corresponding rubric) help the supervisor quantify their observations to be used comparatively to other evaluations.

Speech Analytics, on the other hand, has to be told what is meaningful or not meaningful for a given contact center. For the purposes of determining compliance, this may be easy: mandatory phrases containing specific keywords are helpful clues to the Speech Analytics system. For example:

- "Thank you for calling ABC Company. How may I help you?"
- "This call may be recorded."
- "Do I have your permission to access your policy information?"

Detection of these phrases, and minor variations thereof, can easily be detected. This is bread-and-butter for Speech Analytics. Similarly, Speech Analytics can detect clues about sentiment from the customer that provide valuable data about the customer experience.





Positive Experience	Negative Experience
“Thank you!”	“This isn’t really helpful.”
“You’ve been very helpful”	“Can I speak to your supervisor?”

As evaluation criteria become more complex, however, more and more clues must be used to help the Speech Analytics solution determine whether or not a specific criterion has been met. Consider this question:

- Did the agent offer the product protection policy?

The clues for this question might be more varied and may have decision-tree elements that aren’t as easy for the computer to follow. The language used by the agent to make that offer may come in several possible variations – all of which increase the opportunity for false positive matches. We also need to help the Speech Analytics solution answer other questions: Did the call warrant a policy upgrade? Did the customer ask for it without being prompted? It’s not impossible to provide the necessary clues for a Speech Analytics solution to answer this question – but they’re not going to be there as part of the out-of-the-box solution. Without this second-level grooming, the data produced by the Speech Analytics system is going to be suspect. Useful, perhaps, but not accurate.

ROLE OF SPEECH ANALYTICS

Speech Analytics is a very powerful tool, in that it presents contact centers with the opportunity to scan 100% of audio recordings. That's an opportunity that very few contact centers can afford to pass up. All but the smallest contact centers can claim they have a handle on agent performance and customer issues through their managerial interactions with agents and evaluations of some relatively small sample of calls.

Speech Analytics provides contact centers with a technology that can process 100% of calls and search for sentiment and specific speech patterns.

Yet as we've outlined above, Speech Analytics is not to be used casually. It takes a good bit of "grooming" (not to mention system design and selection). And this grooming is not supervisor work, either; it's the work of dedicated analysts. Which brings us to another conclusion: traditional QA focuses on the specific, whereas Speech Analytics focuses on the general. A scored evaluation can tell you how well agent ABC is doing on particular skills, or the average score as compared between Team XYZ vs. Team PDQ. Speech Analytics, on the other hand, is more adept at identifying strategic insights – displaying score trends vs. the calendar, or correlating scores. Speech Analytics is certainly capable of evaluating agent performance – to the extent that the system is well-groomed, and free of noise. But that analysis is removed from the domain of the supervisor, who is the expert on where agents are in terms of training, coaching, skillset, and personality.



THE ROLE OF EVALUATIONS

Traditional QA, while limited in scope, has the advantage of being far more precise. Consider the questions below – Speech Analytics software, which is limited to detecting audible speech patterns (and some basic qualities of sentiment), would struggle to answer these questions with precision:

- Did the agent establish rapport with the customer?
- Did the agent ask probing questions?
- Was the agent persistent in overcoming sales objections?

Using Speech Analytics, a supervisor would be hard pressed to make an evaluation of either a single agent or a group of agents with any reliable precision. Speech Analytics “scoring” is limited to counting the number of occurrences of specific words or phrases. So what formula do you use to detect and count probing questions? Or building rapport? A human, on the other hand, could nail these every time, given the time it takes to do a proper evaluation. And a human can do so according to an organizationally-adapted rubric – what constitutes satisfactory vs. outstanding.

The questions above, it should be noted, are 1) important, and 2) don't suffer from the lack of scope in which traditional QA is limited. That is, if I can answer those questions for one single agent, it's highly likely that these can be considered “characteristic” of that agent for all (or most) of their calls – these are part of their skillset and not likely to vary by a huge amount from call to call. So the increase in precision more than makes up for the sampling disadvantage.



Consider, on the other hand, the kinds of questions that CANNOT be answered by supervisors using traditional QA; that lend themselves extremely well to Speech Analytics:

- What percent of the time do agents follow (portions of) our script?
- How often do agents mention our brand?
- How often do customers express satisfaction on calls?
- How are agent interactions split across call types?

No doubt that a supervisor could answer these questions as accurately as Speech Analytics software. But clearly, the resulting data is only meaningful when the sample size is fairly large – at least 10x what a “normal” traditional QA program would allow. The Speech Analytics software, however, can do this in its sleep.

WHICH TO USE?

Considering the pros and cons of each, is there a hard and fast rule about which to use, when? Or, could they be used in tandem, to complement each other? The obvious answer is that the best case would be to use both, together.

It can be concluded that supervisors are more capable of answering intricate questions, while Speech Analytics is more precise when it comes to more general, easy to categorize questions. But that’s missing the most important point. The advantage of using Speech Analytics technology will always be that it can collect data from ALL calls, and use that data to quickly spot correlations, trends, and anomalies. Yes, customers should think carefully about which questions are answered by Speech Analytics and which questions are answered using traditional QA. But more importantly, customers should think of how that data will be used.



Speech Analytics and traditional QA are complementary. Neither can completely replace the other. Using both gives contact centers a valuable combination of tools and programs.

A good analogy is to think of a car driving on a freeway at night. The car's headlights will reveal, in detail, what's directly in front of the car. But because night vision is less accurate, cars occasionally use high beams to see further out, so they have more time to adjust to new terrain. Traditional QA is like the headlights, whereas Speech Analytics is like the high beams. Speech Analytics should definitely be used to monitor script adherence (which, by the way, frees supervisors to spend more time on more intricate aspects of agent performance), but customers who limit their use of Speech Analytics to such simple tasks will probably not achieve an acceptable ROI. Speech Analytics should be focused primarily on those things that CANNOT be done using traditional QA, which is the true realm of Speech Analytics.

CONCLUSION

Speech Analytics solutions definitely have a place in every contact center environment. With the importance of customer service, the demands placed on agents, and the operational costs of contact centers in mind, Speech Analytics is a technology most contact centers can't afford NOT to use. But a keen understanding of Speech Analytics – how it operates, what are its strengths and weaknesses – is essential to a successful deployment and ROI.

Traditional QA still has an important place in contact centers, too. Supervisors and Quality Managers have a need for accurate assessments of agent performance, to facilitate coaching, to validate training, and to meet quality and service objectives.

Speech Analytics and traditional QA are complementary. Neither can completely replace the other. Using both gives contact centers a valuable combination of tools and programs. Ideally, they balance each other – Speech Analytics can be used to verify customer experience and agent compliance, and to highlight issues that should be monitored and managed using traditional QA.



Measure What Matters

Encore™ from DVSA Analytics is built to help you take millions of recorded customer interactions and turn them into rich business intelligence and insights.

Encore is engineered to integrate into your existing contact center and UC platform.

DVSA Analytics
Workforce Optimization and Engagement
Management

Gain a competitive edge today!

www.DVSAanalytics.com