

SPEECH ANALYTICS

FOR SOUTH AFRICAN CONTACT CENTRES

By Rod Jones

Version 2.0 First published 20 February 2020



Preface

Although certain aspects of speech analytics have been around the contact centre industry since the 70s, it is only in the past decade or so that the technologies have matured and become viable for practical deployment. Recent developments and integration with advanced analytical tools, artificial intelligence (machine learning) and process automation have accelerated the adoption of speech analytics into mainstream customer contact centre technologies. But not without certain limitations and challenges; not the least of which has been the cost of deployment and the complexities of regional languages, dialects and accents.

My interest in speech analytics grew exponentially in recent months (late 2019) when I discovered that a group of South African scientists and software engineers had unveiled the first-ever speech analytics solution built specifically for South African English (and all major accent variations), Afrikaans, Sesotho, isiXhosa and isiZulu and that developments are well underway to introduce several other of the more widely used indigenous languages. A demonstration was all that it took to convince me that speech analytics is about to become ubiquitous in the South African contact centre environment; following similar take-up and growth patterns in North America, the UK, Europe and Australia.

My recent online research and readings on the subject of speech analytics have taken me on an enlightening adventure, traversing literally hundreds of white papers, articles, analysts' opinion pieces and formal research papers. In this eBooklet I have cherry-picked content from my findings and written-up what I think are the most salient points; the features, benefits and the advantages of deploying speech analytics into South African contact centres.

I sincerely wish to ensure that this eBooklet contains sufficient information to be of benefit to operational managers, IT and support specialists, customer experience managers and executives, and our good friends, the decision-makers in Finance.

To this end, the contents of this booklet have been peer-reviewed by several contact centre industry professionals and other subject matter experts. I recognise, thank and acknowledge them for their valuable contributions in the reference section at the end of this booklet.

Rod Jones

Lonehill. Johannesburg. South Africa. February 2020

What is Speech Analytics?

The more sophisticated of these types of speech analytics solutions also have the ability to identify or detect small variations in speech patterns; shifts in emotions, agent or customer stress, talk-over, abusive callers or the use of inappropriate language, by callers or by agents.

Speech analytics is a software solution that automates the process of 'listening' to and analysing the content of recorded customer interactions – for example, telephone calls. These technologies achieve this by utilising recordings of conversations or interactions between contact centre agents and customers and converting recorded voice to text. Speech-to-text is performed using among others proprietary acoustic models, language models and pronunciation dictionaries. Post-processing of the text (or lattices) can then identify and extract or tag relevant and vital content in the form of individual words or phrases. In addition, these systems are not only able to transcribe and identify what was said (or was not said or omitted from the conversation) but also (in the hands of a skilled analyst) able to interpret the actual substance or meaning of specific parts of the interaction, and in some cases, even the sentiments associated with individual calls.

The more sophisticated of these types of speech analytics solutions also have the ability to identify or detect small variations in speech patterns; shifts in emotions, agent or customer stress, talkover, abusive callers or the use of inappropriate language, by callers or by agents.

** For more technical information relating to speech analytics technologies please refer to the last section in this booklet.

Speech Analytics in South Africa

The global take-up and deployment of speech analytics solutions is expanding at an estimated CAGR of 21%+, with market value burgeoning from US\$1.1 billion in 2018, and expected to reach US\$3.6 billion by 2024. However, utilisation of these powerful technologies by South African contact centre operators has been particularly slow. But this is about to change rapidly!

In a nutshell, the more advanced speech analytics solutions use a technology called LVCSR, or *Large Vocabulary Continuous Speech Recognition* for the analysis of keywords or for phrase spotting. This technical approach is dependent on having specific language and acoustic models and pronunciation dictionaries to be able to correctly identify and transcribe words and complete sentences. To be reliable and accurate, these language models need to take into account all manner of local or regional accents, variations in pronunciation and regional specific words and phrases.

In addition, these language models and pronunciation dictionaries need to be maintained and continuously refined to include, for example, organisation-specific words such as product names, company or organisation names, or industry specific terms such as those used in the medical services sector.

South African Languages have been a major problem



Until recently, the vast majority of speech analytics solutions used in the South African contact centre environment have been unable to provide the significantly high levels of transcription accuracy required to be reliable and feasible to deploy on a large-scale. In addition, many of these solutions, being internationally focussed, are simply not cost-effective or economically viable for the South African market. Most of the current solutions are based on Automatic Speech Recognition (ASR) models developed specifically for the North American, UK or Australian markets.

However, the good news is that after decades of research, a team of South African scientists and engineers based in Cape Town and Stellenbosch have developed a speech analytics solution that is achieving far superior transcription accuracy (when compared to international language packs) for the analysis of South African English (with all our major accent variations), as well as for Afrikaans and isiZulu. isiXhosa and other indigenous languages are to be added as the solution develops further in 2020 and beyond.

The practical application of Speech Analytics

With the advent of a locally developed solution, speech analytics has become a highly cost-effective method of driving operational improvements in terms of effectiveness, efficiency and streamlining the overall Customer Experience. These technologies are being widely used to identify positive and negative drivers, trends, strengths, weaknesses within services, products, competitive offerings, the supply chain or any other critical processes, as well as agent and customer behaviours.

The Benefits of Speech Analytics

Speech analytics tools are also being used extensively to mine vast volumes of recorded information to access valuable, credible and verifiable insights for business intelligence and to drive tactical and strategic improvements in contact centres, and often, throughout the entire enterprise.

Speech Analytics Generates Business Intelligence

Speech analytics can provide many attractive and valuable business benefits to most customer-centric organisations. When appropriately deployed the outcomes and business intelligence that speech analytics delivers rapidly become vital, not only for the contact centre, but also to the organisation.

Speech analytics can deliver significant benefits in three primary focus areas; **Direct Cost Savings** (driven by Operational Efficiencies), **Revenue Generation** (including all aspects of Customer Experience, Sales, Loyalty, Customer Retention and Lifetime Value etc) and **Compliance**.

In addition, speech analytics can be used by management to fully understand what drives many of the typical agent-related contact centre KPIs. As an example: Tracking, monitoring and reporting on agent adherence to scripts or other compliance-related directives. Armed with this knowledge and deep insight, all manner of interventions can be implemented to drive business results or KPIs towards desirable goals and outcomes.

Proven Benefits and Advantages of AI-Based Speech Analytics

Before expanding in considerably more detail the specific benefits derived from the use of contemporary AI-based speech analytics, it is well worth highlighting a few of the widely recognised and acknowledged headline statistics.

Typically, organisational users of speech analytics adept with the appropriate skills, are experiencing significant measurable successes. The following are a few of the typical gains that are being reported in respected industry publications:

- Reduction in call Average Handling Time (15% to 25%)
- Reduction in Customer Churn (Depending on vertical) (10% to 15%)
- Increase in First Contact Resolution (30%+)
- Increase in Employee Satisfaction (30%+)
- Increase in Customer Satisfaction (30%+)

In the case of Collections and Sales, a minimum increase of 10% can be expected. Similarly, cost reductions in the region of 10% to 15% can be realistically anticipated. This solution enables quick and easy deployment, requires very little integration groundwork and is highly cost-effective.

Several of the benefits discussed below have **Cost-Saving, Revenue Generation, Efficiency Improvement** and **Compliance** characteristics.

Cost Savings

The reality of today's economic climate continues to put immense pressure on contact centres to significantly cut operating expenses, but without diminishing customer service levels or reducing revenue aspirations.

Significant cost reductions will be achieved by the appropriate deployment and on-going refinement of the usage of speech analytics technologies. This will unquestionably lead to improvements in operational efficiencies, and consequently, to significantly reduced operation costs, improvement of the customer experience and to enhanced employee engagement and staff satisfaction.



For many organisations, speech analytics will be either something completely foreign to existing operations, or a technology that may have been 'dabbled in', but without producing significantly impressive results. Major factors that tend to contribute to a rather 'ho-hum' attitude to the use of speech analytics in South Africa are firstly, the relatively low levels of accuracy (i.e. High Word Error Rate or WER) of currently available (international) speech-to-text solutions, and secondly, the perceived high costs associated with implementation, integration and usage licencing.

Cost Savings – A High-Level View

Contact Babel (www.contactbabel.com), the globally respected, UK-based contact centre industry research analysts, highlight eight primary cost-saving benefits directly related to appropriate utilisation of speech analytics technologies.

- Reduction in call volumes after understanding and addressing the reasons for customer calls, such as acting to optimize any broken processes elsewhere in the organization (e.g. website, marketing, distribution, etc.) causing calls.
- Reduction in cost of unnecessary call-backs after improving first-call resolution rates.
- Avoidance of live calls that can be handled better by IVR or website self-service.
- Reduction in headcount from automation of call monitoring and compliance checking.
- Avoidance of fines, penalties and damages for non-compliance in heavily regulated industries or certain BPO contracts.
- Reduced cost of Quality Assurance and monitoring.
- Lower cost-per-call through shortened handle times (AHT) and fewer transfers.
- Lower new staff attrition rates and recruitment costs through early identification of specific training requirements.

Increase Revenues

For most commercial contact centre operators (both Captive and BPO), revenue generation is a core feature of the business model. Whilst customer experience, loyalty, retention and satisfaction are equally important, particularly when evaluating new or untried technologies, revenue generation rises to the fore in terms of measurable priorities.

Sales through Service

Much of this incredibly valuable business intelligence can be gathered by analysing massive numbers of past calls, or alternatively, in near real-time. Based on representative and credible data provided by speech analytics, appropriate strategies, tactics, training, coaching, refined or re-engineered processes and other operational interventions can be implemented with high levels of confidence.

Most astute contact centre executives will agree that virtually every customer interaction is an opportunity to generate incremental revenues or to preserve or enhance the prospect of ensuring long-term, sustainable revenues through loyalty and retention. Using speech analytics to build a library of words and phrases that identify positive or successful customer interactions (or failed interactions), provides the organisation with powerful insights and strong guidelines for the development and implementation of refined, revenue-orientated customer engagement strategies, and practices.



Promoting Customer Loyalty & Retention

For many organisations, the cost of acquiring new customers can be as high as eight to ten times the cost of retaining existing customers. It follows that often customer retention can mean the difference between profit or loss.

Speech analytics provides management with extremely valuable insights that can dramatically reduce the number of customers at risk of defecting, even in real-time, by accurately predicting whether a customer's expressed dissatisfaction will lead to churn. This is achieved by detecting trends based on the occurrence of specific keywords and phrases, as well as specific customer requests, competitive brand or product mentions, and even detectable shifts in customers' emotions.

Customer Retention: A Use Case: Mobile Phone Sector

For organisations operating in diverse sectors and industries, competition for market share has become increasingly more volatile. Customers have become savvier and many will 'shop around' to find the best or most suitable deal to meet their specific needs.

In the mobile phone sector, for example, customers have become increasingly wary of long-term contracts, so-called "free handset deals", complex bundles or packages and other short-term promotional offers. They will often challenge their existing provider to meet or better a competitive offering. Most often, these demands will be made via the contact centre.

Speech analytics (either post-call or near real-time) will enable the organisation to rapidly identify certain 'trigger' words and phrases that indicate the threat of churn and the severity of such threats. These insights enable the organisation to craft and refine appropriate counteroffers or deal packages and to train agents how better to respond to these types of customers to achieve higher rates of retention and customer profitability.

Identify Upsell/Cross-Sell Opportunities

Consider that a typical inbound contact centre agent can potentially handle between 18,000 and 25,000 customer interactions a year. Depending on the type of contact centre operation, this represents a huge opportunity to generate significant revenues by means of Upsell/Cross-sell.

Analysis of calls handled by top-performing agents will reveal the most effective techniques that lead to successful sales, to conversions or in the case of collections, higher promise-to-pay outcomes.

This information also guides product or service selection, development of specific value propositions, scripting, training, coaching and mentoring programmes to further enhance sales revenues. Similarly, speech analytics can assist analysts to quickly identify lost or missed sales opportunities or outright rejection of the offer by certain types of customers and identifiable circumstances.



The Voice of the Customer

Mainly due to having limited financial and human resources, many contact centres are only able to pay lip service to the concept of Voice of the Customer (VoC) programmes, claiming that the operation gathers enough customer insights from both unsolicited and solicited sources. Solicited sources typically being post-call IVR questionnaires and unsolicited sources being the insights extracted from call monitoring and call assessments made during the quality assurance process.

Typically, the average contact centre will only monitor and carry out thorough assessments of between 2% and 4% of all calls. Yet we need to ask: What far-reaching strategic and tactical decisions are being based on this minuscule level of sampling?

Particularly in sales-focused contact centres, using speech analytics to analyse historical calls to identify both buying and rejection 'triggers', will provide management with powerful strategic and tactical guidelines in respect to product or service offerings (relative to specific customer segments or call or interaction types).

In addition, speech analytics will enable management to accurately track, monitor and report on agents' adherence to well-proven scripts or techniques or whether or not agents are offering customers certain products or services.

How different it is when speech analytics is used to monitor and assess 100% of all calls in and out of the contact centre; both from what the caller is saying and, of equal importance, what the agent is saying. (Or NOT saying!)

Technology is now able to provide reliable, credible and statistically valid customer insights relating to overall customer satisfaction (CSat), Net Promoter Scores (NPS) and customer effort scores (CES).

Speech analytics allows the operation to drill down into the background data that gives rise to these scores; to truly understand the key drivers and triggers, both positive and negative, and in many cases, to identify root causes of customer dissatisfaction, of broken processes or product or service failures. This is the true value of a comprehensive Voice of the Customer Programme.

Customer Sentiments

It is widely recognised that depending on the type of operation, well over 40% of callers will quite possibly be annoyed or even angry at the moment that their call is answered in the contact centre. Many of these callers will be braced and ready for an argument! But what do we really know about these customers and how to handle them appropriately?

In the case where real-time (or near real-time) speech analytics is deployed, the technologies can identify keywords, phrases and even the acoustic qualities of both sides of the conversation to determine when a call is deteriorating.

Automatically triggered by these in-call events, screen prompts (Screen Pops) can be used to guide the agent to use alternative language, tone or style or to offer alternatives to callers, in order to recover the relationship and to deliver a significantly improved customer experience.

Improve First Call Resolution (FCR)

There is very little debate that for the majority of contact centres, First Contact Resolution (FCR) is possibly the most watched and the most important operational metric. Improving the FCR rate by providing fast, satisfactory solutions to customers' problems or queries within their first call, is a key contact centre performance metric for two specific reasons. Firstly, each subsequent call significantly increases the cost of providing the service or solution, and secondly, additional calls will have a substantial influence on lowering customer satisfaction and consequently, reduced loyalty and spend. Not to mention negatively impacting service level!

Repeat calls are a reality for most contact centres. Speech analytics goes a long, long way towards helping analysts to ascertain the real root cause of low FCR rates, as well as flaws in processes, systems, policies, procedures, supply chain or even the organisation's products, services or value proposition.

The reality is that every repeat call diminishes customer satisfaction and loyalty and erodes the contact centre's operational capacity. Increased staffing in order to maintain service level will increase overall operational costs.

Various case studies clearly indicate that organisations incorporating speech analytics as part of their FCR and customer experience strategies have achieved operational improvements as high as 15%. In addition, simply the awareness that their contact centre is utilising speech analytics can drive positive changes in agent behaviour and a reduction in call avoidance, escalations or the necessity for call-backs.



Improve Average Handle Time

It is widely recognised that AHT is one of the key metrics (from an operational or production perspective) for contact centre capacity planning. In a nutshell, AHT tells us precisely the amount of time an agent is working on or is committed to an identifiable task or work item.

Based on this data, we have the ability to accurately plan for the capacity to meet or surpass customer service expectations and maintain the operation's prescribed service level and to drive improved productivity.

AHT is also one of the most commonly used metrics for overall contact centre operational performance, as well as for individual agent or team performance. The fundamental problem with using AHT as a performance metric is that in many cases it forces agents to focus on speed of service rather than the quality of service. Also, AHT metrics tell us virtually nothing about call quality or call outcomes. There is a growing trend for management to closely monitor AHT whilst agents' actual performance is measured in terms of a combination of contact outcomes and customer satisfaction.

AHT statistics and data on their own provide little in the way of actual insights that will allow management to truly understand the composition of calls. On the one hand, it is known that long calls add costs to the operation. But are these long calls resulting in increased sales, increased customer loyalty and satisfaction, or is poor agent attitudes, training, lack of knowledge, lack of access to information or poor performance at the root of many typical customer experience problems?

Segment Call Types Based on Speech Analytics

By segmenting calls based on the insights from speech analytics, managers are able to redefine the use of AHT-based performance metrics and to focus on actual call drivers and call outcomes, rather than a generic one-size-fits-all approach. This allows agents to provide additional time to service clients who need it, without being negatively impacted.

Similarly, a proliferation of short AHTs may well indicate that certain agents are a lot more concerned with fielding calls more rapidly and are therefore not focusing their time on handling the interactions correctly, or to fully resolve customers' issues.

When used appropriately, speech analytics can guide analysts to accurately categorise each type of call or interaction, and through root-cause analysis, determine what a reasonable length of time might be for each interaction type. Analytics can also be used to compare current AHTs to those that have taken place historically. Such an analysis will provide management with valuable insights and understanding of operational challenges, and opportunities.

Benefits can include: Call type categorisation, differentiated AHT targets, design of appropriate agent training programmes, granting agents access to information and knowledge created to address frequently asked questions (FAQs).

There are literally dozens of other practical interventions that can be crafted to significantly reduce AHT, whilst simultaneously improving the overall customer experience, and addressing operational requirements for efficiencies and effectiveness.



Watch out for Agent Shenanigans!

Speech analytics goes a long way towards supporting contact centre managers in taking precautions and actions to mitigate some typical agent 'tricks' that can seriously skew AHT and other operational performance stats. For example: Comparing call duration and non-talk time with the wrap-up or disposition code of calls can be very informative. An analyst can quickly and easily identify calls of - say - longer than a minute with minimal non-talk time which has been incorrectly wrapped/dispositioned as 'no connect' or 'voicemail'. Using the information in this way for coaching or, for that matter, for disciplinarys, can be extremely valuable as a means to drive desirable agent behaviour and performance.

Bear in mind that the speech analytics solution will be able to accurately identify specific call types or categories and add this to the call metadata to significantly improve the accuracy of the analytics.

The following are more examples of agent behaviours that can be identified using speech analytics:

- Deliberately cutting off long or complex calls as the call duration approaches the AHT target
- Unnecessarily transferring or escalating long, difficult or complex calls
- Not probing for additional customer issues after the primary issue has been successfully addressed
- Not truly listening to caller's problem, challenges or information requests
- Rushing callers to complete calls
- Making promises to callers that cannot be kept
- Blaming or finding loopholes in IT
- Avoiding or not completing after-call wrap-up processes
- Not following up on undertakings to customers



Contact centres are coming under increasing scrutiny and pressure from regulators and other bodies to ensure maintenance and compliance with appropriate regulations, standards of behaviour, specific processes and prescribed practices. Certain operations will have additional responsibilities to maintain strict compliance with clearly defined practices or face dire consequences. These consequences may range from severe financial penalties to full criminal prosecution and loss of the business.

Speech analytics provides a cost-effective way to provide managers with appropriate insights to ensure that agents perform according to these requirements and expectations. As an example: The ability to identify, tag and analyse which specific calls are out of compliance, the reasons for this and the degree of severity of the infringement. Based on these insights the appropriate remedial actions can be taken.

Improve Script Compliance

Especially in highly regulated industries such as banking, insurance, collections and other financial services, call centre agents must meticulously adhere to approved scripts

What's more, speech analytics will also identify when compulsory compliance statements are **not** made or are incomplete. And all this whilst maintaining high operational performance and customer experience metrics.

Using speech analytics, managers can automatically mine or interrogate **all** calls to ensure that agents are following specific approved processes and procedures, and making the correct compliance statements, in the correct sequence within each call.

By utilising insights drawn from speech analytics, analysts are able to methodically refine and make changes to agents' scripts in order, for example, to achieve a better understanding of typical compliance statements by callers, faster AHT and improvements to the overall customer experience.

Very often agents are under huge pressure to meet defined productivity goals or other efficiency-based KPIs. This kind of pressure will often result in a high percentage of non-compliant calls where, for example, agents fail to make certain disclosures or compliance statements. The general practice of assessing randomly selected calls can result in vast numbers of non-compliant calls 'slipping through the cracks' and placing the contact centre and the organisation at huge risk.

Score & Assess Every Call

An appropriate speech analytics solution can automatically score **every** inbound or outbound call against specified compliance criteria. This process can be configured to take place in almost real-time or at least very soon after each call has been completed, with appropriate notifications to alert agents or their respective supervisors of 'non-compliance', for poor performance or for high compliance risk.

Improving Agent Performance

Most contact centre consultants will agree that depending on productivity, as high as 75% of a typical contact centre's total operating overhead can be attributed to staffing costs. With agent attrition rates in some centres reaching 35% and higher, the sheer cost of recruiting, inducting, training and coaching to the point of viable competency, is a daunting prospect. These costs are not only the measurable direct financial costs but also the cost to the organisation in terms of reduced customer satisfaction when customers engage with inexperienced and perhaps under-trained agents.

Utilising speech analytics, 100% call monitoring allows team leaders or supervisors to easily categorise calls by type and to quickly identify poor or unacceptable agent behaviours; to have true insights and therefore the ability to intervene with appropriate, targeted correctional coaching, training and other interventions. This obviously applies to both new recruits and to agents with long tenure.



Quality Monitoring & Quality Optimisation

As opposed to monitoring for risk and compliance, it is a vital element of contact centre operations to monitor customer interactions and to assess these in terms of the operation's requirements and definitions of interaction quality.

Quality Assurance plays a vital role in maintaining the organisation's predetermined standards of agent performance, behaviour, discipline and other aspects of the interaction that impacts the customer experience in any way whatsoever. In addition, QA provides team leaders or supervisors with valuable insights for specific agent coaching, skills development, training requirements or when appropriate, disciplinary actions.

Typically, the QA function evaluates small numbers (typically, well under 5% of total call volume) of mostly randomly selected calls and measuring, for example: Did the agent greet the customer properly?; How courteous or professional the agent was during the call?; Did the agent validate and confirm captured key customer data during the interaction?; Was the customer provided with the correct information in an efficient and friendly manner?

In many cases, this 'traditional' approach to QA tends to focus significant resources on identifying lower-performing, weak, or poorly trained agents or agents with bad attitudes. Certain QA assessors can also introduce biases and discrimination by 'cherry-picking' specific calls or targeting specific agents.

By using this data, the operation can develop and deploy appropriate training, top-up training and coaching interventions across the entire estate. The consequences will be significantly improved customer experience, shorter handle times, and measurably improved agent engagement resulting in longer tenure.

Improve the Efficiency of Quality Assessments

Appropriately skilled and experienced quality analysts are an expensive overhead to any contact centre operation. To use QA's to assess randomly selected calls is extremely time-consuming, costly, and inefficient.

It easily takes three to four times the average call duration for a quality analyst to carry out an accurate call assessment, score or rate individual sections of the call, and then complete the process by adding coaching or training notes. For operations requiring that **all** calls are assessed for compliance, the number of assessors and the overall cost of the assessment process can become exorbitant.

Speech analytics allows more efficient use of QA resources by allowing shortlists of likely low scoring calls to be automatically extracted, to dramatically increase efficiencies and reduce costs, allowing QA and compliance teams to focus only on high-risk or poor-quality calls. The coverage of QA is significantly increased, compared to random selection.

Speech analytics can quickly change the paradigm by focusing on driving quality by the majority and not by exception. In other words, using speech analytics' insights based on 100% call assessments, quality teams can concentrate their attention and efforts on uplifting the quality of 80%+ of the entire centre's agents' performance; not merely only concentrating on the lower or 'delinquent' levels.

Make Sampling Theory Work for You

When applied diligently, Sampling Theory can provide the organisation with statistically robust intelligence drawn from a smaller sample than might otherwise be considered. For example: A random sample of 400 or calls will give approximately a 95% confidence level, irrespective of the total number of calls in the pool. But it is a time-consuming and costly manual exercise to manually QA that number of calls. Using speech analytics to fully comprehend the 'inner workings' of such a sample will certainly provide the organisation with reliable insights on which to base tactical and strategic decision-making. But, just imagine how accurate and how relevant insights will be when speech analytics is used to sample **all** calls!



Streamlining Business Processes

At the very heart and soul of any effective, efficient contact centre are the operation's processes. In many cases these processes are not only the drivers for delivering outstanding customer experiences, but they constitute an important element of the organisation's overall value proposition. It, therefore, stands to reason that processes should be clearly defined, exceptionally well documented, and reviewed and re-engineered on a regular basis.

Speech analytics provides operational managers with deep insights into why certain customer interaction types result in significantly longer AHT or lower CSat; insights into the root cause of issues, identification of 'broken' processes as well as providing strong guidelines for the development of new processes designed to improve operational efficiency, improve customer experience or reduce operational costs.

Process Automation (sometimes referred to as Robotic Process Automation or RPA) has at its foundational level, impeccably designed and thoroughly tested processes. Utilising speech analytics to identify less complex and frequently invoked contact centre processes will help guide management to identify and prioritise rational RPA development and deployment plans.

Improvement in Operational KPIs

In many cases, contact centres use as many as twenty (or many more) operational KPIs as a means by which to monitor and understand performance, productivity and effectiveness. Raw data drawn from the ACD and other systems can provide much of the basic statistical information.

But by skilfully using speech analytics, the operation will be able to mine **every call** for vast amounts of incredibly valuable metadata.

This adds considerably to the context of many of the KPIs, and significantly enhances the quality of the business intelligence that the contact centre can utilise to drive its own performance and to give substance to the organisation's strategic and tactical decision-making.

Check and Mend Service Level

The building, managing and maintaining enough infrastructure and appropriate staffing is a fundamental of capacity planning and the maintenance of service level; that is, the ease and ability of customers to access the contact centre resources. Speed to answer, together with first contact resolution are two of the most important contributors towards achieving optimised customer experience.

Analysis of vast numbers of historical and current calls or interaction recordings will highlight many previously 'invisible' insights and provide management with the ability to identify critical areas in operation and agent performance, such as flaws in processes, behavioural issues or weakness in products, services or distribution channels. Interventions designed to counter these types of issues give rise to far more astute capacity planning and consequently, infinitely better control over service level.

Changing Contact Centre Culture

An inefficient and ineffective contact centre is a toxic contact centre, and it has much to do with contact centre culture. Toxic contact centres have exceptionally high staff turnover, poor customer satisfaction metrics, burgeoning costs, minimal executive support and a highly disengaged and unhappy workforce.

In the hands of a skilled and experienced manager, speech analytics can be used to re-mould contact centre culture and to craft far more efficient and effective processes and practices. By identifying and tracking various speech categories, the true voice of the customer will be heard and new or refined processes can be crafted to deliver an optimised customer experience.

Speech analytics enables the operation to develop and introduce new performance measures and change programmes designed to dramatically improve overall customer service, whilst addressing agent and middle management and support team needs. This can be achieved by using analytics focussing on extracting Voice of the Customer data and overlaying this with quality scores, operational KPIs, root cause analysis, call context, customer satisfaction and effort scores and additional metadata drawn from other contact centre systems.

Fast Return on Investment

How speech analytics will directly and measurably benefit individual organisations will naturally vary from one operation to another and will essentially be driven by how cost-saving and revenue generation strategies and tactics are deployed. Suffice to say; with at least one of the currently available South African-developed speech analytics solutions, a surprisingly fast return on investment is possible.

Because speech analytics projects can be implemented quickly, at relatively low cost and without complex integration, results from improvement projects and programmes can be evaluated very close to in real-time. Insights that may previously have taken weeks or months to extract using conventional (human) analysts and tools can be rapidly produced with high credibility, reliability and accuracy. Speech analytics enables contact centre managers to be exceptionally agile and quickly adapting to ever-changing customer wants needs and expectations.



Increasing and Improving Market and Business Intelligence

In today's fast-moving digital economy, organisations need to be incredibly agile to be able to respond to changing market conditions; to the competitive landscape and to socio-economic realities. Agility represents success, or even the survival of the product, brand, service or even the company.



Selecting a Speech Analytics Solution

Analysts are reporting that by mid-2019 over 30% of North American and European (including the UK) 'Best-in-Class' contact centres had already implemented some degree of speech analytics. The trends are indicating that speech analytics in the contact centre customer service industry, is rapidly transitioning from being in the 'early adopter' phase to becoming 'mission-critical'; an ubiquitous feature of any aspiring world-class contact centre.

Astute business decisions rely on a constant stream of reliable, credible business intelligence on which to base strategic and tactical decision-making. Traditionally, the bulk of business intelligence has been gleaned from the marketplace using conventional market research methods, perhaps in combination with customer surveys, focus groups and various ways of testing products or communication in the field. This approach is not only extremely costly, but it is also time-consuming and, very often the findings are presented far too late for the organisation to derive significant benefits.

Any organisation with a modern contact centre and call recording capabilities has access to potentially hundreds of thousands, if not millions of hours of customer conversations. Herein lies the true voice of the customer. Speech analytics is today's well-proven method of rapidly obtaining data that is both quantifiable and vastly less costly than conventional market research.

By comparison with traditional market research, speech analytics has a number of powerful advantages:

As is the case with most new or emerging technologies, it is vital that organisations thoroughly examine the offered solution, not only for inherent features, but more specifically for quantifiable business benefits, and the cost of delivering these. In the South African context, specifically examine high-quality transcriptions of South African languages and dialects, the ease of use and setup and the operational costs. In this section, the key factors that are ideally the framework for astute decision-making are highlighted.

Low-Risk Evaluation.

Unlike many other relatively new contact centre technologies, depending on the vendor, an initial evaluation of a speech analytics solution can be inexpensive with little or no risks. It is also possible to use initial trials to build a compelling business case based on reliable and credible analysis and reporting on the organisations current or historical call recordings.

Multi Language Functionality

Specifically, in the context of the South African customer service and contact centre environment, it is a prerequisite that a speech analytics solution must be able to provide high accuracy for voice-to-text transcription, as a precursor to using speech analytics algorithms interfaced with appropriate language libraries. In other words, the solution must be able to specifically work with South African English (with all accent and pronunciation nuances), and support Afrikaans, isiZulu and isiXhosa, Sesotho and other regional languages.

Specific words, phrases and company or industry-specific terms

The use of 'jargon' and even appropriate slang are a reality for any contact centre. The speech analytics solution must therefore have the ability to absorb and reference to constantly updated word and phrase 'audio dictionaries', specific to the organisation and its customer base.

Transcription Accuracy of the selected solution is a key factor to consider. A target threshold of the accuracy of $\pm 70\%$ should be considered a minimum. Speech recognition accuracy is typically measured by word error rate (WER)^[2]. (Bear in mind that the type of recording technology used, and the actual quality of recordings will have a significant bearing on transcription accuracy.)

Speed to Business Insights is a major factor to consider. A preferred solution will be able to analyse recorded interactions and provide initial insights within hours, minutes or in certain instances, in real-time or very close to real-time.

Integration with Robotic Process Automation tools. When integrated with appropriate process automation tools, (RPA) the speech analytics solution will be able to identify specific keywords, phrases and even subtle variations in callers' (and agents') emotional state to trigger appropriate desktop prompts, alerts, alarms or escalations.

Reporting capabilities of the speech analytics solution is paramount. Given that the solution will have the capability to analyse massive volumes of calls, it is imperative that it is quick and easy to craft queries, develop standard reports and have the ability to access meaningful insights and present these in a credible and attractive form.

Identification of Call Types. Using combinations of technologies that recognise phonemes (the phonetic 'Lego Blocks of Language') as well as *Large Vocabulary Continuous Speech Recognition* (LVCSR) (Specific language 'audio dictionaries') together with powerful proprietary algorithms, a speech analytics solution will be able to rapidly build a reliable 'library' of identifiable call or interaction types.

Allowing the speech analytics engine to automatically encode call type into the recording metadata dramatically increases the accuracy and reliability of traditional Aux codes^[3], and allows for comparison to metadata captured by agents.

Identification of Trends, Patterns and Anomalies is an important feature of any good speech analytics solution. When used to support root-cause analysis methods, Speech analytics will quickly provide the insights that can lead to fast implementation of remedial actions. For example: The development of appropriate customer self-service options or the deployment of automated responses.

Search Capabilities. To derive the full benefit of a speech analytics solution, many organisations will want to not only analyse current voice recordings but also to find insights or trends or other business intelligence from potentially millions of historical calls. To achieve this, the solution must have powerful and accurate search and reporting functionality; the ability to locate and report on call types, use or absence of specific words or phrases, and even fluctuations in tone, pitch, pace and volume of either callers' or agents' voice patterns.

The use of graphical representation reporting, visual displays, dashboards, automated alarms and escalations are all valuable features of today's more advanced speech analytics solutions.

See (and hear) the entire conversation. In many cases, analysts or managers will need to see the transcription of entire conversations or agent-customer interactions. To provide correct context, the accuracy of such transcriptions is vital, as is the ability to listen to specific sections of the conversation and to add tags or flags or assessor's notes.

Security and Data Privacy are important considerations when comparing speech analytics solutions. Cloud-hosted and on-premises solution offerings will have distinctly different features. These need to align with the organisation's policies as well as with certain statutory and regulatory requirements.

Immensely valuable business insights can be derived from analysis of call or interaction types, relative to (for example) average handling time, revenue per call, customer satisfaction score, net promoter score or customer effort score.

Ease of Use. Business users need to be able to 'hit the ground running' with minimal training or on-going dependency on either the vendor or their own internal IT resources. The solution must, therefore, be intuitive to use and have all the basic functionality out-of-the-box, yet allow users to quickly and effortlessly build queries, run reports and expand the value of the tool with degrees of customisation.

Technical Support and Professional Services. Access to specific technical skills relating to the integration of the speech analytics solution with other technologies is vital. In addition, hands-on training of analysts and other users should be an important element of solution selection. Speech analytics solutions by themselves are merely highly sophisticated enablers. It is the knowledge, skills and experience of the individuals tasked with using these tools that will deliver true benefit to the organisation.



Implementing a Speech Analytics Solution

“A speech analytics solution is not a simple matter of ‘plug ‘n Play. There’s a journey up ahead and it can be a long and winding road to reach the desired destination” Rod Jones.

Or, as that doyen of the contact centre and Cx profession, Martin Hill-Wilson puts it so succinctly,

“Know what you're getting into, because there's a journey involved”.

The Right Solution and The Right People

In short, deploying speech analytics should not be taken lightly. If the chosen solution is inappropriate to the needs of the organisation, or if the organisation's project sponsor lacks enthusiasm or the required experience and skills, the project will almost certainly fail.

Speech analytics needs a rare kind of person to fully leverage its potential to deliver meaningful insights and ROI. It needs individuals who can think logically and who can spot and articulate insights that are relevant to the organisation, and who have the guts to present, and then fight for changes and operational improvements to be made.

Low-Risk Evaluation. Tests, Pilot Projects and Proof-of-Concept. Given the precarious realities of the global and regional economy, most organisations will have stringent procurement policies and procedures in place; particularly when it comes to funding new or untried technologies. The astute user will engage with a speech analytics vendor that will provide a low-cost trial or proof-of-concept, business model. After all, procurement decisions will most likely have to be based on a robust business case built on actual reports drawn from the pilot or proof-of-concept of the speech analytics solution. Only once the critical ROI questions have been answered in the affirmative, will the deployment of Speech analytics make any sense whatsoever to the decision-makers.

Cost! The solution must be affordable.

The Right Problem

In the vast majority of cases, the primary reason why any organisations would consider deploying speech analytics will be that somebody, somewhere in the operation has identified and is struggling with a significant, and probably costly problem. One of the keys to ensure long term success of the Speech analytics project will be to use the technology to solve such problems and uncover insights that will have immediate relevance and direct, measurable and long-term sustainable benefits to the organisation. Preferably financial benefits.

Avoid the Big Bang

The successful deployment of speech analytics will start by solving one problem after another in an iterative approach. As the analyst gains skills and experience, the relevance, the importance and the value of the insights gained in the process will become more meaningful to the business. Both the solution and the analyst will gain credibility, as will the project sponsor.

Get More Value from Quality Monitoring

There is no doubt that the implementation of speech analytics into a practical quality assurance program is likely to be a formidable and highly complex task. As mentioned above, some organisations have 'dabbled in' speech analytics technology without actually producing significantly impressive results. Contact centre culture and agent behaviour can be difficult to change and it is challenging to know how to use the results of the speech analytics solution to encourage positive behaviours.

Use the Speech analytics solution for positive change, not as a disciplinary tool

Speech analytics is not a silver bullet on its own. Just like many new technologies, if there's no buy-in from the outset, the kick-back is likely to be monstrous. Use speech analytics to identify root cause or as a tool to coach and improve agent performance, or to unlock process bottlenecks that give rise to agent frustrations or tensions. Show that the operation is supporting and growing its people, not looking for ways to penalise them.

Change Management is an Imperative

To ensure long term success, the organisation needs to invest a significant amount of time and effort on change management and on devising ways to make the insights gathered from the speech analytics tool available, relatable and clear to the agents; this, in such a way that they can have sight of their own performance, as well as how they are performing relative to their peers.



How Does Speech Analytics Technology Work?

In the context of call centres or contact centres, Speech analytics refers to the technology used to automatically analyse recorded calls or interactions between agents and customers. Today, some contemporary speech analytics solutions don't merely identify individual words and phrases. Sophisticated technologies can be used to identify linguistic and semantic elements of recorded conversations and can potentially identify specific sentiments and intent, and even the emotions of either customers or agents.

There are fundamentally two basic approaches to speech recognition that are commonly used in conversational analytics solutions: *Phonetics-based* and *word-based*, (Large Vocabulary Continuous Speech Recognition LVCSR). And as will be seen in this article, the more advanced solutions may utilise a hybrid of these two models, coupled with a variety of mostly proprietary algorithms and elements of Artificial Intelligence (AI).

It all starts with phonemes...

Phonemes are the smallest unit of sound in a word that makes a difference in its pronunciation, as well as its meaning, from another word. Phonemes, or sequences of sounds, are like the 'Lego blocks' of any language. It is important to note that all languages, dialects and regional or cultural variations of languages and common language usage, have their own unique set of phonemes.

The speech analytics process starts by decoding the audio signal using an *Automatic Speech Recognition* (ASR) model. The output of this process can be either a lattice (a network of different possible phonemes or words that were spoken), or a 1-best output (the most likely sequence of phonemes or words that were spoken).

Some of the early speech analytics solutions relied solely on the phonetics approach. However, hybrid *Phonetic-LVCSR* technologies are now more widely deployed, and more recently developed solutions focus only on word-based recognition

The Difference between Speech Analytics and Voice Analytics

Fundamentally, a *Speech Analytics* solution focuses on the actual spoken content of a conversation by transcribing and analysing what was said. In other words, it looks for specific words and phrases and how these relate with one another. Speech analytics places greater importance on what was actually said and produces transcribed data that can constitute tangible evidence.

On the other hand, *Voice Analytics* works by focusing and analysing specific audio patterns, or what is called, the prosodic features of *how* it was said. Prosodic features are features that manifest when we put specific sounds together in connected speech. Prosodic features can include intonation, pace, pitch, pronunciation, stress and rhythm. Based on these features, voice analytics technologies can potentially detect callers' intent, emotions, mood or temperament.

Most modern speech analytics solutions seamlessly integrate Phonetic, LVCSR, Speech Analytics and Voice Analytics.

Accuracy of Voice-to-Text Transcriptions

It will be appreciated that to have tangible value as the basis for analysis, the accuracy of voice-to-text transcriptions should be as high as possible. The accepted unit for measuring the accuracy of a speech analytics solution is referred to as the *Word Error Rate* or WER; the percentage of errors in the final transcription. It is widely accepted that an average Word Error Rate of approximately 30% or below is sufficient on which to base reliable analytics.

What Factors Drive Accuracy? Language Library

One of the primary contributors to transcription accuracy is the composition of the LVCSR acoustic model. It will be appreciated that to use, for example, an American English or UK English acoustic model to assess South African English (with all the nuances of regional and cultural accents), the accuracy of transcriptions will probably be much lower, with a Word Error Rate much higher than 30%. However, when analysed using a South African language acoustic model, specialised pronunciation dictionaries and a good recording solution, word error rates can drop as low as 10% for broadcast news-like speech.

Quality of Voice Recordings

A major contributor to the accuracy of voice-to-text transcriptions is the overall quality of the voice recordings being analysed. Some of the factors that can produce poor recording quality include the following:

- Voice quality over the connection. A poor or inconsistent mobile or broadband connection. Poor VoIP quality due to inadequate bandwidth, poorly managed WAN/Lan network structure or utilisation. (Packet Loss, Latency, Jitter)
- Poor quality agent headsets not incorporating appropriate noise-cancelling technologies to mask background noise. Poorly positioned agent headset microphone.
- Agent's speech. E.g. Volume, tone, pitch, accent and pace.
- Inappropriate recording technologies or methods. When audio compression is used (such as mp3) and mono recordings, the quality degrades considerably. (Best practice for optimum transcription quality is uncompressed Wav Stereo Recording at 8kHz).

Special Thanks & Credits

During the past two months researching, writing and editing this eBook, I have had tremendous support and inputs from several individuals who's time and insights I truly value. Getting ideas and information and perspectives from industry professionals and subject matter experts has been a valuable source of inspiration. In addition, the collective 'beady eyes' have spotted and corrected umpteen typos, many examples of my atrocious spelling and quite a few grammatical 'oopsies'.

My special thanks go to Francois du Plessis, Corey Springett, Ingrid De Leeuw, Simone Wills, Charl van Heerden, Bryan Schneider and Tanya Phillips.

*Rod Jones
Lonehill, Johannesburg.
February 2020*



**"I help Decision Makers
make Good Decisions about Contact Centres"**

 rod@rodjones.co.za

 +27 (0)82-568-9976

 www.rodjones.co.za



IMPORTANT COPYRIGHT NOTICE

This document is protected by South African and International legislation and common law pertaining to the protection of copyright and intellectual capital contained herein. Rod Jones Intellectual Capital Enterprises (Pty)Ltd reserve all rights and titles and interests in all data, analyses, recommendations, proposals, reports, materials, its knowledge, tools, software, ideas, concepts, methodologies and processes, including, but not limited to market research, information processing, product development, services development, and related strategies which either party uses, acquires or develops for the full term as agreed upon by the parties. Such rights are vested with the authors Rod Jones Intellectual Capital Enterprises (Pty) Ltd. No part of this document may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without the express written permission of the authors.