calabrio

There's no end to better.

Calabrio ONE™

# Speech Analytics - The Power of Simplicity

## A Guide to Phonetics-based Speech Analytics

## Introduction

Effective action requires knowledge. Speech Analytics (also known as audio mining) is a valuable tool for optimizing contact center operations by providing the contact center - and the enterprise - with greater knowledge about what customers are saying and how it impacts the business. With Speech Analytics, contact centers can pinpoint customer satisfaction and performance opportunities by automating the process of "listening" to every call - looking for key words or phrases that may indicate circumstances, issues or trends that require action. This overview examines the power of Speech Analytics technology to extend resources and expand knowledge. It explores the preferred architecture for businesses that want to leverage the capabilities on conservative investments, and defines the steps to a successful start out of the gate – as well as a path to build on that success as the benefits of the technology take hold in the business.

## Key Benefits & Use Cases

Speech Analytics flags calls for review and analysis, which provides a way to mine for key calls of interest more thoroughly AND with less effort. Without human intervention, Speech Analytics lets a business mine all calls to pinpoint business issues or trends that require action. For example:

- Quality issues
- Performance evaluations
- Process issues
- Product or service issues
- Customer defection threats
- Competitive threats
- Compliance breeches
- Business trends

Without Speech Analytics, it simply isn't practical for businesses to apply the resources necessary to find - among thousands of calls - those critical issues that can have a dramatic impact on the business. With Speech Analytics, calls are more easily uncovered for analysis, so action can be taken to protect and improve the business.

## Technology Approaches

There are many technology approaches to Speech Analytics, but two are most prevalent:

- LVCSR (Large Vocabulary Continuous Speech Recognition), which is a derived transcript (aka speech-to-text). In LVCSR a conversion of the call is made into data (text) to be analyzed.

- Phonetics, which looks for defined sounds or strings of sounds within spoken words. In Phonetics, a conversion of the call is never converted, but stays as an audio file to be searched, indexed for rapid search and analyzed.

These two approaches define the underlying speech engine layer. Both approaches require capturing the call and saving the recorded files for analysis – and both have their inherent strengths and weaknesses. Following is an analysis from DMG regarding these two approaches.

The Phonetics approach provides clear advantages for business that are interested in a robust technology that is easy to deploy and manage without an extensive staff of expert resources. For clear applications where the information sought is known (competitive names, escalation issues, general complaints, compliance, etc.), a Phonetics approach allows significantly more call coverage – literally 100% of all calls – due to it's efficiency. Typically 1 hour of audio recording needs only 45 seconds of processing on an average CPU workstation.

To be clear and fair, there are some applications that simply may require the LVCSR approach (e.g., a full test file of a recordings,) however, Phonetics offers a straight-forward and effective alternative for a wide range of applications that deliver a significant benefit to the bottom line of the business, (e.g., customer satisfaction and retention, revenue and compliance.)

*Continued on next page*

## Best Practices for a Phonetics-based Approach

***A powerfully simple alternative for Speech Analytics.***

A Phonetics-based engine provides users with the ability to search large audio files for keywords and phrases with dramatically increased speed and accuracy. A Phonetics engine searches the way a word sounds, not the way it is spelled, allowing it to recognize uncommon words such as names, places or acronyms, even if you don't know the way that it is spelled.

Speech is one key advantage. A Phonetics engine can index and process speech up to 80 times faster than real-time. The Phonetics engine also maintains an "open" vocabulary at pre-processing time, allowing searches for any word/phrase without having to re-process the audio. This makes the Phonetics-based approach significantly faster than alternative LVCSR methods, which require extensive time and processing power to transcribe and analyze speech.

Another great advantage of a Phonetics engine is its ability to adapt to changing requirements. Want to add a new project? Or even a new language? A Phonetic engine can do it faster, at far less cost.

***How does the engine work?***

**Phase 1 - Indexing (Pre-processing)**
A Phonetic representation of the audio recording is created in real-time and is stored as an index file. The indexing phase is a once-only analysis of the speech file, performed in order to optimize the speed of subsequent searches. This pre-processing or indexing can be done on a routine basis and as calls are captured and stored.

**Phase 2 - Query Parsing**
Query parsing is a once-only analysis of a user defined search string or phrase producing a Phonetic representation of those words or phrases of specific interest. E.g. "cancel my order," transfer me now" or "excellent service."

**Phase 3 - Searching**
The search phase utilizes the index data to search for the words or phrases of specific interest. The Phonetics engine will find all of the places where it is likely that the given word(s) or phrase(s) were spoken. Each such "hit" will have a start and end time to identify the exact location in the audio recording where the word(s) or phrase(s) occurred.

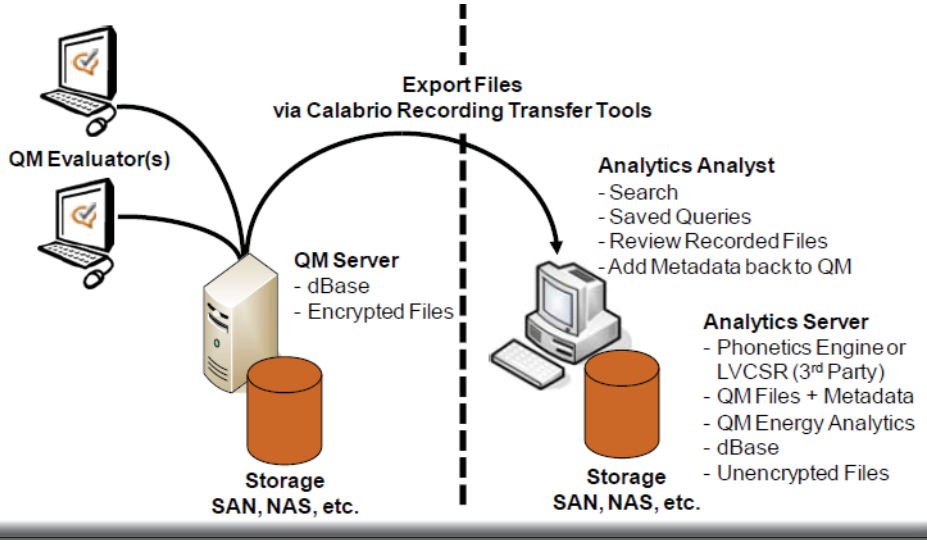
## Architecture for Success

The best architecture for analytics will be a result of the analysis of the business problem to be solved and the analytics technology and timing to be applied. Following is a recommended architecture and model for supporting Speech Analytics with minimal complexity:

***Analytics Services***

One architectural approach is to separate the analytics on a standalone workstation from the recording infrastructure, which provides a low risk method for analytics experimentation as an initial deployment or trial system. (See the following, Figure 1, for a high level view on deploying this system.) Alternatively, a similar approach as described can be integrated within a quality monitoring and evaluation system, which is beneficial if the output of the analysis will be tightly tied-in with quality evaluations.

QM & Speech Analytics Architecture




[ Figure 1 ]

For the purposes of our discussion, we'll describe the stand-alone approach. The Analytics workstation will receive unencrypted audio recording files and associated metadata from the recording infrastructure and then apply the analytics technology to these file copies. This separates the analytics processing load from the recording system and prevents any possibility of impacting the recording system's playback capabilities. The Analytics workstation also hosts the analytics query and the playback application that will be utilized by Analytics analysts to meet the business objective. The Analytics workstation will provide the following key functions:

- Convert and process recorded files from their native capture format(s). Selectively export recordings by configurable policy to balance the business process need with the available processing and storage capacity of the Analytics workstation.
- Search by word, phrases, energy level and metadata for "Calls of Interest."

***Tuning for Accuracy***

There are trade-offs to be made between thoroughness and time. To ensure you never miss a call, a contact center would set the confidence threshold low. This means you flag more calls that are less likely to be hits to ensure that you listen to more calls. Set the confidence threshold high to flag only those calls that the engine has a high level of confidence as a true hit. This will be driven, by some extent, by application. If you are merely looking for triggers and indications of service issues, you may want to tune the engine down to hit only those calls with a high level of confidence. If the application is directly tied to revenue or compliance/risk, you may want to tune the engine up to ensure no opportunity or threat is missed.

***Metadata Adds Context***

It is important in speech analytics to not only analyze what is being said (words or phrases – Phonetics), but the data around the call to gain a clearer picture of events. This data, know as metadata, can be either CTI, external data like CRM account information or energy based type data. Some of the key data elements supported by Calabrio are:

- Original called number
- Original calling number
- Date of call
- Time of call
- Duration of voice call
- Recording reason code – one of longest, shortest, first, last or random
- Agent name and ID
- Team name
- Number of instances that the agent placed the caller on hold
- Duration of the longest hold
- Total amount of time on hold
- Speech energy - silence and talk-over events

Adding metadata to the query is a powerful way to key-in on a particular type of call or particular attributes of the call for analysis, saving time and increasing the effectiveness of an analytics program.

## Conclusion

Speech and Content Analytics are powerful technologies for improving customer satisfaction and contact center performance. But, you cannot realize the full benefits of Speech Analytics until the problem to be addressed is defined. Ultimately, the key to success is defining and executing against a clear business goal - understanding the business problem(s) an enterprise needs to solve:

**Step 1** - Determine business needs and problems or areas you are looking to analyze for speech or content analytics. Start with clear and well defined objectives with straightforward queries that help you learn and understand the tools and technology as you realize the benefits – uncover competitive threats, look for instances of escalation or frustration, and find calls that are out of compliance.

**Step 2** – Assign the proper resources for the project. Someone within your organization must be assigned the task of finding, analyzing and defining actions against key calls of interest.

**Step 3** – Determine the speech analysis technology (LVCSR, Phonetics or both) which is best suited to solve the business problem. Phonetics is an excellent choice for companies that have a clear understanding of calls they want to find and conservative resources to support the project.

**Step 4** – Set up a system and use search queries and contact review to refine search and process criteria for implementation into the business best practices.

**Step 5** – Measure success and build from there!

Post call analysis can identify processes, improvements, trends and personnel issues to help improve future operations, but it will not save today's lost sales, service problems or missed opportunities. So much about a business can be learned by listening to customers – and where better to focus those efforts than in the contact center? By starting slowly and building on success, companies are sure to find that Speech Analytics is an invaluable tool for growing and improving the business.