Speech Analytics: It’s all About Statistics

By Bob Moore, Aspect Software
Speech Analytics: It’s all About Statistics

By Bob Moore, Aspect Software

SPEECH ANALYTICS is one of the most ground-breaking technologies to appear on the call center scene in recent years. As DMG Consulting puts it, “Speech analytics is a true ‘change agent’” with a unique ability to efficiently convert unstructured voice information into structured data that is highly reflective of the voice of the customer. However, given the power of speech analytics to open new paths to intelligence about the customer and the operation, combined with the increasing speed of adoption, best practices for taking maximum advantage of the technology continue to evolve—quickly.

When you consider implementing a speech analytics solution, there is a fundamental decision to be made before all else: Which of the two major approaches to speech analytics should you choose... phonetics or LVCSR (large vocabulary continuous speech recognition)? Both phonetics and LVCSR approaches process incoming speech by identifying phonemes (unique elemental sequences of sounds), and then identifying words, phrases and even sentences that were spoken. This is, by definition, a statistical matching process, which means that sometimes the matches are wrong, introducing errors into the remainder of the process, whether phonetic or LVCSR. The differences between the two approaches are all about how they reduce or eliminate those matching errors.

In the case of a phonetic speech analytics solution, the sequence of phonemes is then stored in a data structure that is optimized for searching on phoneme sequences, a “phonetic index.” When you want to search for an occurrence of a word/phrase, that word/phrase is also translated into phonemes, and the best statistical match(es) with the phonetic index are retrieved. Note that the phonetic indexing and search methodology does not assemble the phoneme sequences into words until search time. With LVCSR, the phoneme sequence is matched against a large dictionary of words containing their phonetic representation and serially translated into text—at the time of original indexing. This is also a statistical process, since phonemes may not match exactly. Searches are performed by indexing through the transcript of translated text. As with any process whose goal is the structuring of unstructured data, these processes, by their very nature, will be incorrect some of the time. You only need to try Siri to know that this is not a perfected science.

Both approaches will retrieve recordings that do not contain the desired word/phrase, and both approaches will miss recordings that do contain the desired word/phrase.

Here is the critical difference between the two. In the case of the LVCSR solution, the audio has already been translated to text at the time of your search. When you search on “wireless,” you get only those audio clips that these error-prone approaches identified as “wireless.” You might actually be getting clips of someone saying, “worthless” or “we’re less,” and you will be missing clips from people with unusual accents. In the case of the phonetic search, the system has not translated the phonemes to something as rigid as a textual word that is either matched or not. The phonemes continue to exist as part of the search process, so you can set a “confidence threshold” at the time of your search. You can set it to low confidence if you want to get all of the occurrences of a word/phrase, or you can set it to high confidence if you want to get only the occurrences of a word/phrase.

This may seem like a fine point, but speech analytics is used for widely varying purposes in the contact center. For compliance verification, you want high recall. You need to be absolutely sure you get every occurrence of “my lawyer,” even if you recall some that need to be thrown out; e.g., “tri-layer” or “fly higher.” For targeted listening purposes, you want high precision because you are not so concerned that you don’t get every instance of a word/phrase, but do get enough that you understand the underlying dynamics. So if you search for “cable service,” you do not mind missing some instances as long as you get a representative sample.

LVCSR is a better solution for applications like transcribing a television show for closed captioning. However, if you have a large number of speakers with different accents, poor audio quality and conversational unscripted speech, and you need to find events, phonetics is a superior solution.

BOB MOORE is senior manager, product marketing for Aspect Software and leads product marketing for the workforce optimization suite, including workforce management, quality management, performance management and analytics. With over 20 years of experience in strategic planning, product management and marketing for technology companies, Bob has helped large and small companies achieve their growth objectives.

PHONE: (978) 250-7900 (North America)  
WEB: www.aspect.com  ■ EMAIL: info@aspect.com
As the omni-channel customer experience becomes the new gold standard, customers expect high quality interactions in all channels of communication.

Aspect® Analytics for Speech and Text lets you clearly understand the voice of your customer in all customer channels. Monitor trends and gain valuable new insights in seconds rather than hours using 100% of your customer interaction data. Use these new insights to create real-time agent desktop guidance and supervisor alerts.

Learn how Aspect analytics solutions help you create remarkable customer experiences for your omni-channel customers.

www.aspect.com
About Contact Center Pipeline

Contact Center Pipeline is a monthly instructional journal focused on driving business success through effective contact center direction and decisions. Each issue contains informative articles, case studies, best practices, research and coverage of trends that impact the customer experience. Our writers and contributors are well-known industry experts with a unique understanding of how to optimize resources and maximize the value the organization provides to its customers.

To learn more, visit: www.contactcenterpipeline.com

This issue is available online at: ContactCenterPipeline.com