

Operations Practice

# From speech to insights: The value of the human voice

In a digital age, live voice contacts matter even more in providing a high-quality customer experience. New technologies make voice-data analysis easier to achieve, for lasting results.

*This article is a collaborative effort by Yasmine Belkacemi, Eric Buesing, Arpit Goenka, Vinay Gupta, Damian Lewandowski, and Maurice Obeid, reflecting views from the Operations Practice.*



© Tom Werner/Getty Images

**The evolution of** the traditional call center into an omnichannel contact center has allowed companies to view the function less as a cost driver and more as an opportunity to provide strategic, experience-oriented customer care. With customers engaged via SMS, websites, chats, and social media, identifying customers' reasons for initiating contact has become a core analytics use case for virtually any contact-center operation.

This increased focus on customer care requires today's leaders to manage more data than ever before—and as more transactions migrate from in-person channels, data management becomes even more important to the customer experience. But many businesses still struggle to capture and process customers' voice conversations. Across industries, these interactions represent the majority of all incoming volume, and projections suggest that these calls aren't going away anytime soon. An inability to analyze voice conversations makes it difficult to unlock the full potential of digital investments and analytics to drive significant customer-service improvements.

Organizations traditionally face three primary challenges as they attempt to understand the direct voice of the customer. First, they follow random, manual call-sampling methods, which capture less than 2 percent of all interactions, producing incomplete (or unrepresentative) raw data sets. Second, organizations work with legacy processing systems that transcribe speech to text, but poor accuracy seriously limits how much useful information can be extracted. Third, organizations try to turn accurately transcribed conversations into meaningful insights, which too often fail to yield measurable initiatives with bottom-line impact.

However, as digital tools continue to improve, natural-language-processing capabilities—paired with industry expertise—are helping businesses improve quality, efficiency, and customer experience. Speech data offer customer insights that simply aren't available from other sources, helping to identify the causes of customer dissatisfaction and revealing opportunities to

improve compliance, operational efficiency, and agent performance.

The results include cost savings of between 20 and 30 percent, customer-satisfaction-score improvements of 10 percent or more, and stronger sales as well. Companies that fail to leverage this information risk falling behind their peers as speech analysis becomes a fundamental expectation across contact centers.

## **The potential of better recognition and analysis**

Improving call-center performance can be a source of frustration for many businesses because of the roadblocks to extracting and understanding voice data. These complications take many forms, but speech analytics frequently offers a solution.

For example, a specialized transportation company depended on voice calls to book, cancel, and modify service orders, but didn't know what proportion of calls fell into each service type. This created multiple issues: poor forecasting, over- and understaffing, and imprecise coaching—the last because the company wasn't clear on what skills needed support. Speech analytics helped the managers understand the reasons customers called, and the company's staffing and team training improved dramatically.

At another company, customers complained about needlessly long handle times: more than 60 percent of calls included more than 20 seconds of continuous silence. The company had historically blamed the problem on limited agent knowledge, but in-depth call reviews revealed that the problem was primarily due to slow systems, as well as a lack of standard procedures and practices across the call center. Speech analytics helped the company quantify those issues and prepare a business case for a systems upgrade.

A third firm's average handle time (AHT) was consistently 10 percent above target. The company had given all agents the same training, regardless

of each agent's performance, instead of offering targeted coaching on individual improvement needs. Speech analytics helped the organization uncover each agent's strong and weak areas, which enabled the company to reduce its AHT by providing each agent a more specific training prescription.

In some cases, the underlying problems turn out to be surprisingly simple. One hotel chain discovered that the sound quality of its call centers' recordings was too low for analytics tools to work. Limitations in processing algorithms and audio quality meant that agent and customer voices were recorded on a single line, making it difficult to differentiate their voices. A combination of improved diarization and an upgrade from mono to stereo recording helped to distinguish speakers, effectively unlocking the data inside the recorded conversations.

### AI advances offer improved speech recognition

It is now possible to solve these and many related difficulties in extracting and using call-center

data, specifically unstructured voice data. Academics and other researchers have continued to push the study of human speech, using advanced artificial intelligence and natural-language-processing models and technologies to power advanced computational linguistics. The newest approaches to automated speech recognition use neural-network language models that take more data into account, enabling more accurate transcriptions. Many analytics providers have strengthened their offerings with complementary capabilities, including security features such as automatic data masking and password protection. These tools may incorporate options for deployment either on-premises or via the cloud, depending on a company's infrastructure and data-hosting strategy (see sidebar, "Questions to ask a speech-technology provider").

A few pitfalls nevertheless persist. But through awareness, careful planning, and judicious intervention, companies can overcome them with a few broad strategies.

## Questions to ask a speech-technology provider

**With many providers** offering speech technologies, companies need to understand each offering's capabilities and how it would integrate with existing operating systems, IT platforms, and strategic priorities.

**Language models.** How accurate is the solution in transcribing speech to text? Which languages and dialects does the system support? Does the solution include machine learning for continuous improvement, or is it strictly rule based? What methods does the provider use to create new language models, and how much data do those processes require?

**Natural-language processing.** To what degree can the solution extract intent and sentiment from a conversation? Can it recognize entities and correctly identify relationships between entities, concepts, and utterances?

**Quality assurance.** Does the system measure transcription quality? Does it guide users to provide ongoing coaching?

**Analytics and technical capabilities.** Does this solution recognize emotion, or offer insights on callers and agents, such as demographics and personality? How

customizable is the platform? How easy is it to use? How scalable is it?

**Real time.** Does this system analyze conversations as they are happening, or is there a lag between conversation and analysis?

**Use cases.** What out-of-the-box use cases does the system provide? How easy is it to customize these use cases to the organization's industry or business model?

# Combining speech data with other customer or telephony data shows the full context of a call, which is often crucial to its meaning.

**Unclear or timid use cases.** Some companies, unaware of the overall value of speech analytics, can't imagine what they would do with data derived from speech recognition. Consequently, analytical teams may generate basic insights into customer sentiments without direction as to how best to use the findings—for example, whether to aim for reduced call volumes, increased sales, or improved customer satisfaction.

**Poor contextual recognition.** A manual, word-based tagging approach to understanding what customers and employees intend in a conversation often leads to poor categorization. For example, a customer might say, "I am not very happy with your service." If a data dictionary captures only "very happy" from that sentence, it will miscategorize customer intent. Or the approach might recognize just one of the multiple intents present in the same interaction, as in the sentence, "I received the product two weeks late and it was faulty." Brand recognition can also be complex, particularly when a product or company is known by a nickname or abbreviation, in addition to a standard form.

**Limited analytical capabilities.** Speech recognition is less valuable when it isn't integrated with other data sources. Combining speech data with other customer or telephony data shows the full context of a call, which is often crucial to its meaning. For example, if a customer has called multiple times about the same issue, then the root cause of the customer's frustration might be related to poor processes, rather than to a particular agent's ability

to solve the issue. An organization can only know that the same customer has called multiple times about a single issue by leveraging customer data.

**Limited scalability.** At some organizations, use-case implementation is the responsibility of only the analytical team. Without buy-in from other business and functional units, however, progress can quickly falter, particularly if the voice-analytics changes have to compete for scarce resources among many technology and strategic priorities. Impact may also fall well short of expectations if units fail to use the resulting insights because they weren't involved in their development.

## Natural-language processing and analytics

Applications of speech-analytics capabilities can yield endless use cases, from sales to operational excellence, and can be tailored to specific industries. The following is our list of standard, nonexhaustive cross-industry use cases.

**Increase data coverage.** Traditionally, organizations sample incoming calls for quality an average of two to four times per agent per month. Examining all of the available unstructured voice data, rather than just a sample, can sharpen the insights generated.

**Monitor KPIs.** A personalized data-visualization dashboard lets clients see any number of conversational moments, from supervisor

escalations and compliance violations to customer satisfaction and AHT. This can help organizations see how well their implementations have gone and measure change over time.

**Accelerate time to insights.** Automated AI transcription drives faster analysis and full call coverage. These moves can accelerate traditional diagnostics time by nearly 400 percent, helping organizations implement recommendations much faster.

**Uncover hidden inefficiencies.** By monitoring a variety of contact-center KPIs, organizations can unearth inefficiencies and identify root causes while hearing the true voice of the customer.

**Personalize training.** In addition to tracking operational KPIs, building organizational and cultural elements into the foundation of performance management is another important task. Agents are crucial actors in helping convey what the customer is experiencing during customer-service calls. With deep insights on every customer call an agent handles, leaders can create custom coaching sessions for individual agents and supervisors, raising customer-satisfaction levels.

**Improve customer experience.** Sentiment analysis lets teams look at the factors driving positive customer engagement, such as empathy statements, and indicators of negative customer experiences, such as supervisor escalations.

**Create better interactions.** Action-based insights generated by speech analytics can create a more positive external environment for customers to communicate with the brand. It also informs management what tech-enabled strategic initiatives will lead to the highest returns on investments, and how the company should prioritize these initiatives.

**Uncover automation opportunities.** Speech analytics can reveal automation opportunities. For example, extended periods of silence during calls are often an indicator of automation opportunities.

**Improve self-serve options.** Speech analytics can indicate the percentage of unsuccessful self-serve calls, break down those unsuccessful calls by category, and determine the percentage of calls in which the agent educated the caller about self-serve options. Companies can use the resulting insights to improve specific self-serve options that many callers have found problematic.

**Increase upselling and cross-selling.** Insights gathered from speech analysis can indicate which agents frequently succeed in upselling and cross-selling. Agents who do not succeed as often can receive specific coaching around these skills.

## Capturing call-center data

Companies that successfully capture, extract, analyze, and act on call-center voice data get a better sense of why their customers call, develop insight into how they can provide a better experience for both their customers and employees, and measure the customer satisfaction of each interaction. This is typically a huge improvement over using infrequent and unreliable customer-survey data. Capturing call-center data starts with five steps.

**Create a list of use cases.** These could range from targeted coaching, automated quality assurance, understanding customer sentiments, managing the workforce, reducing fraud, reducing collections, or increasing sales. One internal help desk provider defined ten use cases that helped it unlock 20 to 30 percent cost savings and a customer-service improvement of more than 10 percent.

**Think systematically.** Align use cases with the industry and strategic objectives. Metrics are simply data points, but organizations can think about the entire call-center operation, rather than a single score.

**Partner with the right speech-analysis provider.** Many vendors can consult on recording sound quality, provide raw transcriptions, offer an easy way to understand intention from transcriptions, identify

emotions correctly, and generate granular metadata to decompose the call.

**Listen to the front line.** One company set up a cross-functional team among QA, analytics, and business-unit leaders to test speech insights aggressively with frontline staff. Another decomposed handle time into multiple mini segments, which helped supervisors effectively coach agents (exhibit). A third broke down the interval call volume into granular call and subcall types, giving the center a much better estimate of its true workload.

**Show me the money.** Getting the most of any analytic capability requires linking measurement and initiatives to financial value. That connection

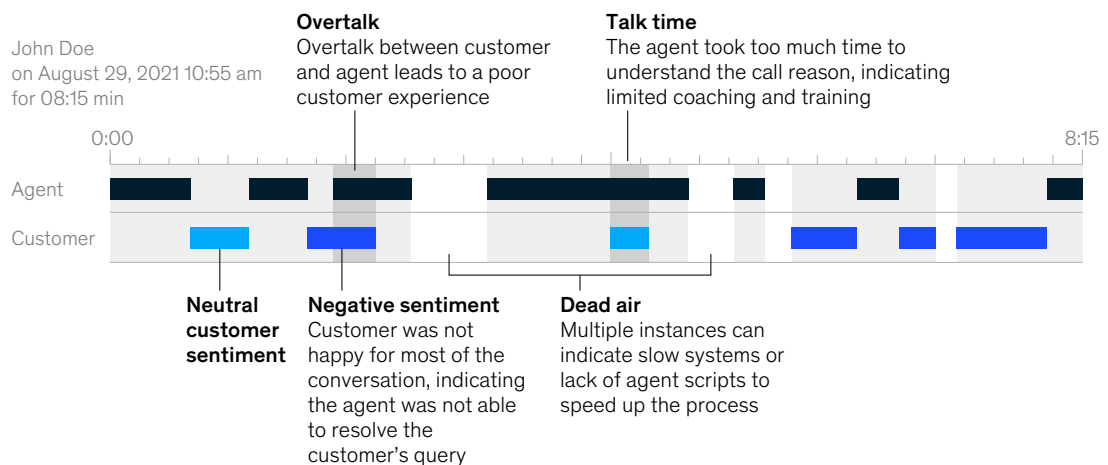
can help business leaders set priorities on the changes that will deliver the greatest bottom-line impact and stimulate improvements at all levels of the contact center.

Generating actionable insights continues to be challenging. Sophisticated listening tools can provide valuable data, but organizations need expertise and insight to understand the implications of that data and use them to drive measurable initiatives with bottom-line impact. Investing in effective and complete analytics capabilities that interpret voice-analytics results, measure the experience, and implement changes rapidly is the way to reap the rewards.

Exhibit

## Detailed voice analysis reveals opportunities for substantial improvement in customer experience.

### Illustrative diagram of voice analysis and insights gained



**Yasmine Belkacemi** is a consultant in McKinsey's New York office, where **Damian Lewandowski** is a new capabilities manager and **Maurice Obeid** is a partner; **Eric Buesing** is a partner in the Stamford office; **Arpit Goenka** is a knowledge specialist in the Gurugram office; and **Vinay Gupta** is a senior knowledge expert in the Waltham, Massachusetts, office.

Designed by McKinsey Global Publishing  
Copyright © 2022 McKinsey & Company. All rights reserved.