# Best Practices for Voice Biometrics in the Enterprise:

## Simple Secure Authentication across Mobile, IVR, and Contact Centers

»

opusresearch Report

opusresearch Report »

# Best Practices for Voice Biometrics in the Enterprise:

## Simple Secure Authentication across Mobile, IVR, and Contact Centers

**Lower costs, happier customers, and faster authentication are just a few of the benefits of voice biometrics authentication that over 150 enterprises have discovered after successful deployments.**

January 2015

Dan Miller, Senior Analyst, Opus Research

**Opus Research, Inc.**
**350 Brannan St., Suite 340**
**San Francisco, CA 94107**
**www.opusresearch.net**

# Table of Contents

## Introduction: Learnings from Hundreds of Implementations

With the global population of enrolled voiceprints over 60 million and successful deployments spanning nearly every continent and vertical industry, it is time to take stock of the business practices and technical approaches that correlate with successful deployments of voice biometrics.

This document is the result of structured interviews and follow-up emails with dozens of solutions providers, system integrators and enterprise executives who have successfully deployed voice biometrics-based technologies as part of their customer care or internal authentication and access management fabrics. It is designed as a practical reference guide for executives in companies who are evaluating their options for customer or employee authentication and looking for insights based on real-world experience with voice biometrics.

Opus Research acknowledges and thanks executives from the following firms for providing input. In alphabetical order, they are: Agnitio, Auraya, Flare Design, Nuance, OneVault, SESTEK, Speechouse and VoiceTrust, with additional input gleaned from interviews and presentations by executives from firms who have shared experience from their successful implementations at Opus Research's Voice Biometrics Conferences or during webinars.

The findings are compiled in a way that takes into account the implications of implementing new techologies and business processes across multiple business units and customer touch-points, including those in charge of customer care, contact centers and mobile offerings. Findings also take into account the vital interests of Security, IT, Finance, Marketing and all business units that take part in evaluating and procuring infrastructure hardware and software.

## Getting Started: Playing the Inside Game

Decisions around authentication and fraud prevention in contact centers often involve teams of managers from customer care, security, contact center operations and IT. The introduction of voice biometric-based technologies is seldom regarded as the top priority of these executives. But our research has revealed that a well-managed rollout of multifactor authentication methods that includes voice biometrics helps companies achieve and exceed their goals for measurable improvements in customer (and employee) satisfaction, as well as reduction in call duration and fraud loss.

Real-world experience teaches that, in most instances, using voice as a factor for customer authentication and fraud reduction is a marked improvement over current solutions. Successful implementations are most often the product of

focused efforts by a team of employees with guidance from a project champion armed with a clear understanding of the role of voice in the company's overall identification and verification strategy (ID&V). Specifically, successful deployments find areas where voice is clearly superior to a company's current solutions, be they PIN or password, knowledge-based authentication (KBA) methods or one-time-passwords.

While it is most obvious in the contact center, implementation of voice biometrics-based solutions is a set of processes that can take months to carry out and involves participation by executives that cross several departmental silos. Therefore one of the highest priorities is to engage in educational activities that get everyone from contact center agents to the head of security on the same page, aware of the time that will be involved in implementation and enthusiastic in support of the technology's ability to overcome known challenges and deficiencies of their present systems.

Initiating voice biometrics-based solutions provides an ideal opportunity for implementers to enlist the aid of hundreds of employees and, ideally, a proportion of its clients and customers.

### Allow Sufficient Time for Tuning

If a text-dependent approach is taken – meaning that customers are required to enroll by repeating a specific passphrase three times – implementers must build a "background model" that is trained to "hear" that passphrase. Over time, the model, itself, will become more accurate as more individuals are enrolled and more authentication events transpire. It is a form of machine learning that can take into account the aging of customers and their voiceprints and changes in communications channel characteristics. There are different considerations when taking a "text-independent" (or "free speech") approach, as we will address below.

It is important to note that the process of building domain-, company- and application-specific reference models can be done in parallel with other activities associated with system initiation and deployment. It can be built on a separate system and put "into production" when the time is right. Allocate at least six weeks for collecting utterances and turning them over to speech scientists for them to work their magic.

> ALLOCATE AT LEAST SIX WEEKS FOR COLLECTING UTTERANCES AND TURNING THEM OVER TO SPEECH SCIENTISTS FOR THEM TO WORK THEIR MAGIC.

While it may be tempting to employ a phone bank or build a special cubicle for individuals to call into the system, this has been characterized as a "worst practice." In the ideal, calls will originate from a mix of phones and environments that represent the variety of devices and background noises that will be common for the community of users. That objective is overcome if all of the calls originate from the same device in an acoustically controlled room.

The need for a rich and diverse sampling of utterances is especially important in instances where your company will be deploying a "text-independent" approach to building a voiceprint and for authentication. In this context, "text independence" means that a speaker can be identified or authenticated during the course of a normal conversation. Enrollment and training based on a specific passphrase is not required. Instead the voice biometric engine builds a voiceprint by capturing and distilling over 20 seconds of continuous speech. Then it can return a confidence score by seeing how well that voiceprint matches with utterances captured in the course of a conversation with an agent or company representative. Even though a specific passphrase is not used, the reference model should include utterances in the local dialect using terms that are customarily used in the course of carrying out phone-based business.

Both text-dependent and text-independent voice biometrics engines will work “out-of-the-box” to a certain degree. Nonetheless, in every instance, performance is improved and thresholds are established by building a “background model” that takes into account the common quality of communications lines, the mix of channels employed (wired versus wireless), language differences and the like. Building the background model relies on capturing utterances from at least 400 people, ideally using the words that they would normally use in the course of their commercial conversations and originating from their accustomed devices and noise conditions.

### Choose Proper Performance Metrics and Achievable Goals

When an enterprise starts its investigation of voice biometrics vendors they tend to focus on the “accuracy” of the core engine. Can it achieve “zero false accepts”? What are the trade-offs in terms of “false rejection,” meaning that a legitimate customer is turned away because he or she cannot be authenticated? Based on vendor-supplied information, today’s text-dependent voice biometric engines can achieve “sub 1%” false accepts (letting the bad guy in) with a corresponding false reject rate of less than 3%.

Lab-based accuracy figures always come with some qualifiers, akin to “your mileage may vary.” That’s because the accuracy of any particular authentication engine is determined by a number of factors that are independent of the core engine. These include line quality, background noise, channel differences (e.g. enrolled over wireless/authenticated over a traditional phone) and others. For purposes of comparison, use the pilot mentioned above as an opportunity to set expectations or “thresholds” that are acceptable and defensible throughout the organization.

For most companies, the precise number of false accepts and false rejects achieved in production are a well-kept secret. It is our understanding that a 90% “success” rate (meaning that an individual is able to authenticate his or her claimed identity) should be considered acceptable. As we discuss later in this document, voice alone is seldom if ever the only factor used to authenticate. Companies customarily know the origin of the call, the device being used, time-of-day, location and other data in order to gauge risk that should be associated with a particular transaction and the importance of adding more layers of security to the authentication process.

> ONE OF THE VERY HAPPY DISCOVERIES OF OUR RESEARCH IS THAT CUSTOMER SATISFACTION HAS A POSITIVE CORRELATION WITH EMPLOYEE SATISFACTION AND RETENTION.

### Better Yet, Work Toward Measurable Business Outcomes

One of the very happy discoveries of our research is that customer satisfaction has a positive correlation with employee satisfaction and retention. In effect, making the authentication process shorter and less loathsome creates a work environment for agents that is much more pleasant. This translates into better productivity, lower training costs and longer tenured therefore more experienced employees.

## Successful Enrollments At Large Scale

When gauging the successful rollout, the first order and most visible metric is “enrolled users.” The most successful implementations (for example, Turkish telecommunications giant Turkcell) set ambitious goals for enrolling new users as they went from zero to 13 million in the space of 18 months. It is a metric that generates pride and accomplishment throughout the company and its partners.

Success rates for authentication is another closely monitored metric. As mentioned above, 90% can be considered acceptable. This term refers to instances where the interactive voice response (IVR) system prompts users to say their passphrase so that it can be compared to the enrolled voiceprint. With a successful rate of 90%, it shortens the amount of time it takes for a customer to accomplish his or her desired task and increases the IVR containment rates. It also reduces the amount of time a customer service agent spends, measured in terms of average handle time (AHT) in order to authenticate the individual and begin to serve his or her needs.

AHT is once again shortened when "passive" methods are used to compare spoken words to a stored voiceprint because agents are not required to pose challenge questions of the callers. Companies are witnessing higher customer satisfaction levels and loyalty, measured through the classic net promoter score (NPS) because customers are not required to do anything but speak in order to be authenticated by the text-independent system.

## Pilot if Necessary, but Keep it Short Sweet and Relevant

In addition to building the background model, our investigation showed that small-scale deployments (or pilots) can also be used to build prove the performance assumptions that are the foundation of the ROI assumptions that justify the procurement. If the solution targets a reduction in fraud loss, the pilot can provide hard-dollar evidence of reduced incidence of fraud. It will also provide hard evidence of the savings resulting from significant reduction in average handling time of calls. In large companies, where seconds in agent engagement amount to millions in average savings, the hard-dollar savings will, likewise, be self-evident.

Because they can drive up implementation expenses, pilots are considered a "necessary evil" and both implementers and vendors recommend that the pilot be based on hardware, software and interfaces that can easily be "moved into production" when it comes time to launch the service. This approach keeps start-up and development costs down and fosters a speedier ROI.

## Choose an Appropriate and Effective Passphrase

When taking a text-dependent approach to voice-based authentication, users are most often prompted to enroll by repeating a passphrase, such as "At [your business name here], my voice is my password." As alternatives, they may be told to repeat the series of numbers from "0" to "9," asked to repeat their account number, or they may be permitted to use a passphrase of their choice.

In terms of best practices, voice scientists counsel against user selected passphrases, while security experts suggest that saying account numbers or phone numbers aloud pose a set of obvious security issues by making nearby people aware of what might otherwise be an unknown identifier. Five-to-seven seconds of speech is required to support high confidence levels in authentication. As it turns out "my voice is my password" falls short of that threshold, which is why the addition of a company name or "At my bank…" has become common practice.

In instances where individuals are prompted to repeat "0" through "9," those recorded digits can be used in a form a liveness testing. An authentication application serve the digits up in random order and display them on a screen or smartphone, requiring the individual to speak the sequence of digits within a seconds or fail to authenticate.

Based on real-world experience and feedback from customers, some businesses have opted to leave their brands out of the authentication process. Their customers have told them that the passphrase can feel like an endorsement and that it feels strange to use it when calling with a complaint or an urgent issue to resolve.

## Promotion, Education and Market Conditioning

Successful launches rely on concerted communications and educational efforts, both internal and external in nature. Their objective is to show the general public the advantages of using voice authentication and recruiting them to enroll their voiceprints to take advantage of the new technology. As we discuss later in this document, the communications strategy to support rollout may also include marketing and promotional messages offered in the form of TV commercials, direct mail marketing, bill stuffers and dialogs with live agents.

Note that your solution vendor can provide a package of services to assist in the start up and market conditioning phase. Those with years of experience and millions of enrollments are well-positioned to assist in assuring the success of your voice biometrics-based initiatives.

### Establish a Communications and Education Strategy

Conducting and administering a pilot or controlled rollout provides an opportunity to launch a communications program that explains the overall objective of deploying voice-based authentication and manages expectations surrounding the sequence and timing of the major steps toward broad deployment. In many cases, the professional services organization associated with the vendor, or the system integrator responsible for installing the platform and launching the services will be able to assist with the messaging and targeted delivery.

Early establishment of a communications strategy will be very relevant as the deployment moves from pilot to production. Much of the content, including the overall objective of the deployment, description of the technologies, instructions for enrolling and getting started and solicitation of user feedback, can be used to communicate with the broader set of end-users. Incidentally, the feedback loop mentioned above is a vital component of the communications strategy. Implementers must listen carefully to end-users and react quickly to any perceived objections, pitfalls or speed bumps. One example of an effective feedback loop is designing a voice biometric authentication FAQ page on a company's website.

> BECAUSE HUMANS RESIST CHANGE, EVEN SUCCESSFUL COMPANIES HAVE REPORTED A MILD BACKLASH AGAINST BEING FORCED TO ENROLL.

### "Opt Out, But Value the Customer's Time and Convenience"

Enrollment rates are dramatically different for companies that adopt an "opt-out" approach to recruiting new participants. Empirical evidence (based on over a decade of deployments) shows that success rates for enrollments reach a maximum of around 35% when customers or clients are invited to use the service (meaning "opt-in"). By contrast, that number will reach 80% success when customers are enrolled through a process that requires them to decline expressly (or "opt out) of the registration process.

The high end of success rates result from messaging that encourages customers to "take advantage of a new techologies to speed up the authentication process." This messaging can be delivered in an IVR prompt at the beginning of an inbound call. After a customer has called in and authenticated in the traditional way (probably entering a PIN), he or she is told that the company is offering a better and speedier way to get to a representative and carry out business.

Because humans resist change, even successful companies have reported a mild backlash against being forced to enroll. In almost every case, they have found it necessary to take steps to soften negative reaction, primarily by promoting the advantages of biometric-based authentication. Their key message has been that callers can always "say no" and not create a voice print but, by doing so, they will miss out on the opportunity to save time in future calls. They also play up the fact that individuals will no longer have to remember the answer to challenge questions or adhere to onerous rules about changing and managing new, obscure passwords.

### Supporting Media to Managing Customer Expectations

A combination of e-mail marketing, billing inserts, TV ads and agent assistance to inform customers of the need to create a voiceprint to support speedy authentication in the future have been brought to bear to build customer excitement and foster enrollments. In many cases, solutions vendors will offer a set of "launch services," which can include promotional materials as well as a team of service professionals that create educational and promotional materials to explain the voice authentication to the masses, including what it is, what they can expect when they register and (later, authenticate), why it is secure, how it works and how they benefit.

Another best practice is to use the introduction of "a new way to authenticate" as a way to demonstrate techological leadership. For example, an eastern European bank used such messaging to foster its innovative image.

Customer segmentation plays an important role in messaging strategy. Inside an enterprise, the executive suite can require all employees to register their voiceprints as a condition of employment, in order to support password reset or a spoken token for access to the company's virtual private network. In a similar way, government pension funds can make voice biometric-based authentication a condition "proving life" and receiving payments.

In most instances, businesses don't have the luxury of requiring their customers to enroll. Indeed, they may not want to offer the service to all of their customers and may opt to select a target market that will (a) benefit from the service and (b) offer the company a positive lifetime value over time. Many of the first implementations of "passive" authentication have been launched with a relatively small group of "high net worth" clients who were engaged in conversation with their financial advisors.

### Establishing Ground Truth

A key part of enrollment is making sure that the individual who is providing a voiceprint is, indeed, the person that he or she claims to be. This is called "establishing ground truth" in industry parlance. For many years, it was regarded as a key challenge, especially for applications, such as government sponsored "proof-of-life" offerings where the people being enrolled have no previous relationship with the business or government agency. In such cases, face-to-face appearances are required and the enrollee must bring a number of documents (such as a passport, drivers license or birth certificate) to give the enrolling party confidence that the claimed identity is valid.

Fortunately, in most of the mass market implementations that we have taken into account, enrollees have a prior relationship with the companies. Those individuals can use an existing credential, like their existing PIN or Password, to establish that they are who they claim to be. Then they can be prompted by an IVR, customer service representative, or financial advisor (as the case may be) to engage in the enrollment process.

Bear in mind that enrolling a voiceprint as a credential for future authentication is very much like changing a password. When it is carried out remotely, the business that is accepting the enrollment has a number of channels that it can use to ensure that the person is who he or she claims to be. We're all familiar with the routine. They may send an email with a link to a web site that confirms authenticity. They can do the same thing with a text message and embedded link. In other words, all of the known methods for discouraging imposters can be employed at the time of enrollment.

Another technique for preventing false registrations is to involve customer services representatives or advisors in the enrollment process as part of a natural conversational flow. Once an agent has gained confidence that an individual is genuine, he or she can describe the steps it takes to enroll (i.e. "repeat a passphrase three times"), including the anticipated time it takes. The agent can also describe the advantages of enrolling in terms of saving time in the future without sacrificing security.

## "Passive Enrollment" is Different, And Growing in Popularity

Looking at the history of voice biometric implementations, the vast majority have been "text dependent," thus requiring "active" enrollment by customers or clients. Looking ahead, Opus Research has observed heightened interest in "passive" methods both for enrollment of voiceprints and for using those voiceprints either for detecting impostors in real time or for conversational customer authentication. It is important to note that customer application and use case will dictate the choice of either passive or active system.

### Capturing 40 Seconds of Speech

In order to generate a "strong" voiceprint from the passive capture of spoken conversations, roughly 40 seconds of spoken words (stripped of silences) is required. Once a voiceprint is created, the system can return a "score" that distinguishes a customer from an imposter in roughly 10 seconds. Confidence in the score increases as the conversation progresses. Using financial services settings as an example, where average call times are roughly three minutes, it is easy to capture the raw material in the course of routine conversations.

One of the preconditions is that the contact center's communications or recording systems have separate channels which enable the system to capture the caller's voice, as opposed to the agent. In targeted vertical industries, such as financial services and government, it is routine to capture and store all conversations in order to comply with government requirements. Such call recordings are a good source of raw material for passively creating voiceprints. However, as discussed below, it is important to let customers know that their recorded voice may be used for authentication purposes and in many cases obtain their permission to do so.

### Obtaining "Informed Consent"

Banks, airlines, retailers, insurers and others already preface incoming calls with the familiar "This call may be recorded for quality and training purposes." To comply with "informed consent" requirements, we advise businesses, at a minimum, to change the greeting to include the words, "This call is being recorded for quality, compliance and future security applications."

> YOU SHOULD ALSO SEEK LEGAL COUNSEL TO DETERMINE WHETHER ADDITIONAL MEASURES MUST BE TAKEN TO INFORM CLIENTS AND GAIN CONSENT.

You should also seek legal counsel to determine whether additional measures must be taken to inform clients and gain consent. Alternatives are informational letters, e-mails or billing inserts with descriptions of passive enrollment and subsequent authentication. Lawyers may say that a formal act of consent is required, which may require a further edit of the greeting to say "by proceeding, you accept that your voice will be used on future calls as an alternative to a PIN, password or challenge question."

## A Checklist for the Future: Keep it Simple, But Risk Aware

For readers of this document, your prime objective is to make authentication as simple as possible while applying the proper level of security. No more, no less.

Keeping things simple often means masking the complexity of all the processes that take place "in the background." In Opus Research's "Voice Biometrics Census: Global Tally of Voice Security and Authentication Implementations," we found that 41% of enterprise implementations representing over 75% of enrolled individuals took place through IVRs in contact centers. This indicates that, in spite of an expressed belief that customers might consider the overall enrollment process to be "cumbersome," roughly three-quarters of enrolled individuals use a passphrase. As we have noted earlier, the vast majority of deployments are "text-dependent."

Simplicity (which equates to convenience), in a multi-channel, multi-modal world means that customers are spared from remembering and providing passwords or PINs and that, in any case, they do not have to repeat themselves as they move from Web site to IVR or IVR to agent. But it must be balanced with the need to assure customer security, which calls for application of  technologies and techniques that eliminate fraud.

### Mix Voice with Other Factors, Including Other Biometrics

The use of multiple factors to authenticate customers is not only a best practice, it is a given. Whether customers are using apps on a smartphone, browsing the Web or making an inbound call, the companies that they are dong business with are already aware of many of their attributes. The smartphone serves as a physical token (or something you have) which also provides such data as location, originating phone number and device characteristics that can give a company great confidence that the device and the transaction that is about to be carried out is legitimate.

Biometric factors such as voiceprints, facial recognition, fingerprints or iris scans, are not just a nice-to-have feature, they are the basis of very strong assertion of identity. In the best case,  combining voice biometrics with a passive factor such as device identification can provide a multi-factor but single-step authentication process. The combination of multiple factors has proven to be a very strong crime deterrent.

> THE USE OF MULTIPLE FACTORS TO AUTHENTICATE CUSTOMERS IS NOT ONLY A BEST PRACTICE, IT IS A GIVEN

### Next Up: A Move to The Cloud

Today's best practices may not apply strongly to the future. IT and security departments, who oversee deployment of both customer care and security infrastructures have long been reticent to move customer data "outside the firewall." Therefore, even though the past year has witnessed a rapid migration of enterprise information processing and data management to the cloud, voice-based authentication has been a premises-based phenomenon.

That means that voice-based authentication processes have largely been premises based. But companies should not want to be on the wrong side of history when it comes to deploying voice biometrics. It can no longer be considered "best practice" to keep authentication on premises. Many self-service and contact center applications have already moved to cloud-based resources, making it a multi-billion industry. It is only natural to think that authentication technologies associated with customer care, which needs to be tightly bound with customer workflows, would be in the cloud, as well.

The move to cloud-based resources also opens the door for companies to leverage data and security resources that can be shared across multiple companies. Credit card issuers share (or federate) data to identify stolen cards and prevent fraud, one can expect multiple companies who have knowledge of known fraudsters and, conversely, the ability to identify/authorize valid individuals and transactions. Companies who can authenticate individuals or ban imposters by listening to their voice should be ready to share information in order to reduce fraud in real time.

This is, after all, the age of the sharing economy.

## Finally, Accommodating a “Mobile-First” Strategy

As 2015 approaches, applications that make the most of smartphones, with fingerprint readers, front-facing cameras and multiple communications channels (data, voice Bluetooth) call for enterprises to take a hard look at their mobile strategies. Voice biometrics have a rightful role to play in the course of phone-based commerce. Banks like Tangerine, in Canada, or ING in the Netherlands recognize voice’s value and have integrated voice biometrics-based identification with their mobile assistance.

These apps demonstrate the value of voice biometrics for bolstering security and for promoting highly personal services. This secure mobile assistance defines emerging best practices for both user authentication and transaction authorization over mobile networks.

## About Opus Research

Opus Research is a diversified advisory and analysis firm providing critical insight on software and services that support multimodal customer care. Opus Research is focused on "Conversational Commerce," the merging of intelligent assistant technologies, contact center automation, voice security and authentication, indoor location, enterprise collaboration and mobile commerce. **www.opusresearch.net**

---

**For sales inquires please e-mail info@opusresearch.net or call +1 (415) 904-7666.**

This report shall be used solely for internal information purposes. Reproduction of this report without prior written permission is forbidden. Access to this report is limited to the license terms agreed to originally and any changes must be agreed upon in writing. The information contained herein has been obtained from sources believe to be reliable. However, Opus Research, Inc. accepts no responsibility whatsoever for the content or legality of the report. Opus Research, Inc. disclaims all warranties as to the accuracy, completeness or adequacy of such information. Further, Opus Research, Inc. shall have no liability for errors, omissions or inadequacies in the information contained herein or interpretations thereof. The opinions expressed herein may not necessarily coincide with the opinions and viewpoints of Opus Research, Inc. and are subject to change without notice.