Active vs Passive Voice Biometrics

Authored by SymNex
Sponsored by LumenVox

Executive Summary

When implementing Voice Biometrics, you will need to choose between two methodologies: Active and Passive. Active uses a fixed passphrase in an automated application. Passive authenticates in the background of an agent conversation. Both modalities have their advantages and the impact of each will vary according to the nature of your business. The key factor in determining the right mode for your organization should be the likely level of customer adoption. The customer adoption rate depends upon your registration processes performance. Successful registration depends upon identification and authentication of the caller, the provision of their consent and the technical completion of Voice Biometric enrollment.

This paper objectively lays out both modalities in detail and the full range of considerations to make the best choice for your business use case.

Overview

The Authentication Challenge

Every organization supporting customers remotely needs to ensure that the person they are speaking to is the person they claim to be. The performance of this process can be evaluated against three, often conflicting, dimensions:

- **Security** – How to minimize the likelihood that someone can successfully impersonate the genuine customer?
- **Usability** – How easy is it for the genuine customer to complete the process and what is the impact of that process on their perception of your organization?
- **Efficiency** – How much time and effort are spent by both the caller and agents to complete authentication?

Traditional Authentication

Traditionally, authentication of customers remotely has been knowledge-based; dependent upon something both the customer and the organization know. The process of authenticating relies on a series of challenges from the organization and responses from the customer. This approach raises several key issues:

- **Security** – Knowledge-based authentication's inherent weakness is that both parties need to be able to recall and test the information. This creates the risk that it can be captured and reused by others. Human nature is also such that we are also likely to reuse information between organizations, forget secrets and be easily tricked into giving them out when we shouldn't.
- **Usability** — The customer is often forced to recall obscure facts or secrets that at the very least disrupts their experience in seeking support and at worst is impossible. They are therefore unable to receive the service they need.
- **Efficiency** — As a result of the twin usability and security challenges many organizations have developed a complex multi-step process for the authentication of callers combining secret information with personal and relationship facts. These can be quick and easily automated for experienced (often more frequent) callers. However, in many cases this requires manual intervention by agents and often several minutes of talk time to complete.

Voice Biometrics

Voice Biometrics is a form of inherence-based authentication; dependent upon something the customer is, which is perfectly suited to authenticating customers in the voice channel as it addresses the challenges with traditional knowledge-based authentication. It makes a statistical comparison between the caller's voice and a previous recording known to be the customer. At a high level it improves performance in the key areas:

- **Security** – It is significantly more secure than traditional knowledge-based authentication because a caller's voice is unique to them; like a fingerprint.
- **Usability** — There is nothing for the customer to remember.
- **Efficiency** — It can be automated or take place in parallel to the service interaction with and agent.

You can read more about how Voice Biometrics work in general here.

Key Implementation Questions

While Voice Biometrics improves each of these dimensions, the extent to which it does is dependent upon the context of your organization and the implementation mode you choose, Active or Passive. This paper helps you understand the key factors that you should evaluate before determining the most appropriate mode of operation for your unique context.

Definitions

Active Voice Biometrics

This form of Voice Biometric authentication most often takes the form of a caller repeating a specified passphrase such as “My Voice is My Password” in an automated application. The customer will be asked to record the phrase several times during a registration process. Because the phrase recording and the authentication step require effort from the customer over and above their reason for calling, we refer to it as Active.

This mode may also be referred to as Text-Dependent Voice Biometrics which reflects the way in which the underlying statistical comparison takes place. The algorithm is optimized for the way in which the callers speak the specific words of the passphrase. This approach may also substitute the passphrase with a random digit challenge and response, telephone or account numbers or even customer chosen passphrase. Each approach requires the customer to record the specific words in advance.

Passive Voice Biometrics

Most often this form of authentication takes place in the background of the call, as the caller is speaking to an agent. It is not dependent upon the customer repeating the same phrase every time. This is also referred to as Text-Independent Voice Biometrics to reflect the fact that the underlying algorithm is optimized for the way in which the caller generally speaks. The genuine customer will have registered on a previous call as a result of a conversation with an agent. This call could have been about an entirely different subject matter. There is no additional effort required for the customer other than saying "yes," which is why it is referred to as Passive.

Identification and Authentication

Many organizations conflate these two processes as a single Identification and Verification (ID&V) or similarly named process, so the boundaries between the two are blurred. For the purposes of this paper and Voice Biometrics in general, we need to make a clear distinction:

Identification: We define identification as the process of correctly matching the customer to the caller. This often involves some form of account or personal identifier such as an account number or social security number.

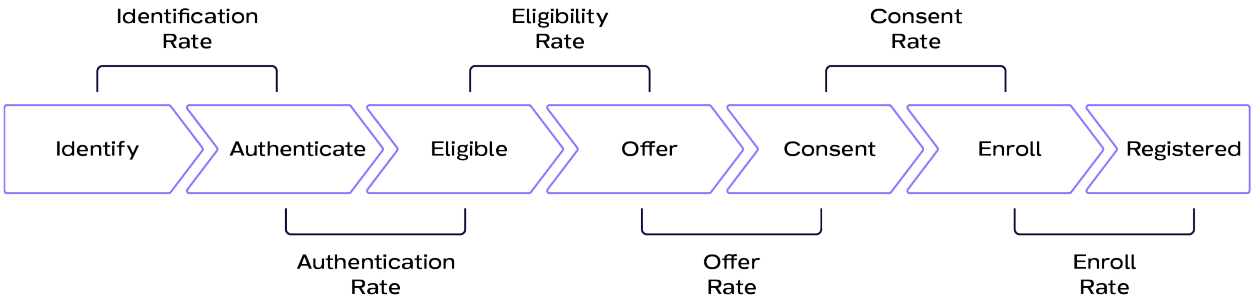
Authentication: We separately define authentication as the processes of testing that the caller is who they claim to be. In the context of this paper Voice Biometrics is used for authentication, so we need to understand the callers claim of identity before the right voiceprint can be tested to confirm whether they are genuine or not.

Adoption

The key factor in determining which mode makes sense for your organization must be the extent of user adoption. All benefits are directly proportionate to the level of adoption, a function of the registration process that establishes customers voiceprints. The characteristics of each mode combined with the context of your organization can make a dramatic difference to the outcomes of this process and overall adoption rate.

Registration Value Chain

We define registration as the set of business processes that ultimately concludes with customers enrolling their voiceprints in a Voice Biometric system so that the organization can subsequently authenticate their identity with it. It includes the following steps:



Identification – Before enrollment we must know which unique record in your system of record their voiceprint relates to.

Authentication – Before enrollment we must ensure sufficient confidence that the customer is who they claim to be and that we can trust subsequent authentications with a voiceprint.

Eligibility – Before enrollment we must determine whether we want to authenticate the customers' identity with a voiceprint in the future.

Offer – Before enrollment we should ask the customer if they want to use Voice Biometrics for subsequent authentication.

Consent – We must get permission from the customer before enrolling them.

Enrollment – Finally, we must create the voiceprint to use on subsequent authentications which for Active authentication may involve the customer repeating the phrase several times.

Please note that in many jurisdictions there are specific privacy requirements that may necessitate more formal disclosures and recording of customer responses for some of these steps.

Active Registration Approaches

There are two principal approaches to registering customers for the use of Active voice biometric systems although in our experience it is generally necessary to implement both to realize the full value of the investment:

- **Automation-Led Registration** – In this approach callers are identified and verified in an automated system using traditional methods and then, subject to some eligibility checks, offered the opportunity to enroll either as a distinct menu option, as a result of calling a specific number or, optimally, in advance of routing to an agent or other automation for their servicing need. If they provide their consent, they are asked to record their passphrase several times before being transferred to an agent or completing other automation tasks.
- **Agent-Led Registration** – In this approach callers routed to an agent who are either identified and verified before routing by an automated system or by the agent are invited to enroll by an agent, usually as a result of a screen prompt that only appears if the customer is eligible. If they provide consent, then the caller is transferred to an automated system at the end of the call to record their passphrase. This system may, depending on regulatory requirements, ask them to confirm their consent before recording their passphrase several times.

Passive Registration

Passive registration uses the same Agent-Led Registration approach as Active but with the distinct difference that there is no additional step required by the customer to complete once the agent has recorded their consent. The only dependency is that sufficient customer audio has been captured on the call. In most cases this can include all audio from the moment the call was connected to the agent.

Active and Passive Adoption Rates

As you might guess from the descriptions above, there is potential for the outcomes of these approaches to vary significantly depending on several factors and how they relate to your organization:

Automated Identification and Authentication Rates

Automated-Led Active Registration approach is entirely dependent upon existing levels of automated authentication. If callers cannot be verified by an automated system, then they cannot be registered without agent intervention. Some organizations are very successful in authenticating customers this way. However, there will always be a number of callers who don't successfully complete these processes because they call too infrequently, or they object to them on principal. For this reason, we recommend that Automation-Led Registration be supplemented with Agent-Led Registration for Active implementations even if only for discrete customer groups. While it is possible to enroll callers before their identity has been verified (which could be done later the same call with an agent who also gets consent or even by subsequent out-of-band communications) these experiences are very hard to make intuitive for callers and have even higher non completion rates so cannot be recommended in most cases.

Offer Rates

We often find that getting agents to offer the service to customers is one of the most challenging aspects of implementation. There is a natural reluctance in many agents that needs to be carefully addressed with appropriate training, incentives and performance management interventions. Even then we still see significant variation between individual agents. Clearly this reluctance does not apply to automated systems but for agents there are ways to inhibit annoyance: the requirement to prevent offers on certain call types, and controls to prevent the same customer being asked too many times.

Consent Rates

In our experience the likelihood of a customer giving consent varies enormously depending on how they are asked and by whom. Automated systems by their nature will only be able to use one of a handful of pre-recorded messages while agents can tailor the initial message more closely to the customer's needs, based on their understanding from the rest of the call, as well address specific concerns or objections they may have. Finally, the perception of effort on behalf of the customer will also influence their likelihood to accept and there is a big

difference between just saying "Yes" and having to be transferred and repeat a phrase several times even if this process often takes less than a minute.

For this reason, we often see significant variation in consent between Automated Active (typically lower than 40%), Agent-Led Active (typically around 60%) and Agent-Led Passive consent (often higher than 80%).

Enrollment Rates

The final step is the completion of the enrollment process itself. For Active, this requires the caller to correctly record a specified passphrase usually three times, but occasionally more if a recording is not deemed good enough by the system. Because this step is automated and time-consuming, we often see lower completion rates than with the Passive registration which is simply dependent upon the call being long enough (which in our experience it is for more than 95% of calls). For callers transferred to Active registration at the end of the call there is a significant proportion who accept the agent offer but subsequently hang up because they have had their service needs met, another group with both agent and automated registration hang up or say nothing because they didn't understand what they agreed to and are put off by the process. Finally, because Active systems are dependent upon the same thing being said, there is a speech recognition requirement to check they have said the right thing that has its own error rates and may reject utterances from the customer.

Costs of Registration

One clear difference between Automated and Agent-Led registration is the latter requires an investment of agent time to make the offer, overcome objections and gain consent. While high performing agents can easily complete this in less than 60 seconds, the impact on average handle time during early implementation--when nearly every caller will be eligible for registration--should not be overlooked. The distinct advantage of combining both approaches for Active is that a significant proportion of callers can be registered without manual intervention.

Other Considerations

As you can see from the above discussion, there is no one way that is best because it is so dependent upon each organization's situation. While adoption is key to successful implementation of Voice Biometrics and should be the principal factor considered, when choosing between the two modes of operation, there are a number of other considerations that organizations should bear in mind:

Efficiency

The efficiency opportunities between the two modes of operation do vary significantly. Voice Biometrics increases contact center efficiency in two ways:

- **Reduced Agent Handle Time** – As a result of removing the need for agents to ask knowledge-based questions both Active and Passive methods can save significant agent handle time.
    - **Active** – For Active almost all authentication time is eliminated. In situations where only automation led registration is used this is generally only substitutional from the previous method so results in minimal net benefit to the organization. When agent led registration is used, we also see more callers authenticate who could previously only identify with an automated system, so we do see a net reduction in agent handle time.
    - **Passive** – For Passive the vast majority of agent handle time is eliminated because there is no authentication effort required. In most cases the authentication result is returned before the caller has even finished explaining their reason for calling. However, we generally see some residual authentication effort, as agents occasionally must fill this time or deal with mismatches.
- **Reduced Agent Call Volume**
    - **Active** – If the automated authentication rate increases as a result of implementing Active voice biometrics, then there is a greater addressable audience for your existing self-service features. Additionally, as a result of biometric authentication requiring significantly less effort on the customer's part we also see a greater propensity to engage (sometimes up to 10%) with these self-service features resulting in increased containment and reduced agent calls even from those customers who have just substituted knowledge-based authentication for biometrics.
    - **Passive** – Today there are few examples of Passive implementations in the IVR, as the technology has only recently reached the level where this is feasible, but we do expect this to be a major part of the opportunity as it introduces the possibility of increasing levels of self-service for customers who previously only spoke to agents.

Usability

There are also usability considerations that vary between modes of implementation:

- **Speed** – While Active authentication takes place in an automated system Passive authentication happens in parallel to the agent conversation so there is a requirement to ensure the result is delivered to the agent in a timely fashion. The process is usually dependent upon a certain duration of the customer's speech (often referred to as "net speech") being provided to the system before the comparison is made. Vendors often quote 10 or 15 seconds as the recommended amount but in reality can return results with high confidence much quicker. In our experience this rarely causes an issue, as callers usually take longer than this to confirm their claim of identity and explain their reason for calling, so should not be a deciding factor.

- **Accuracy** – Historically, even with greater amount of audio discussed above Passive systems have been considered less accurate (i.e. for the same risk of accepting an imposter they reject more genuine callers). While this may be technically correct in theory and laboratory experiments; over the last few years the difference has reduced to a level where it is no longer material to the decision-making process. Both can achieve very low genuine customer rejections in even the most risk averse environments.

- **Cross-Channel** – There is an argument that the short length and integration approach of an Active passphrase makes it an ideal candidate for use in authenticating the same customer in other channels such as face to face, web/mobile applications or even smart speakers. While we don't disagree and there are some good examples of this deployed in mobile apps today, each of these channels has its own unique considerations that are likely to favor different authentication modalities. Our recommendation is always to choose the optimal modality for the channel under consideration and if it can be exploited elsewhere then that is an additional benefit but should not be the principal decision factor.

- **User Perception** – In our experience customers do not perceive significant security differences between Active and Passive but do favor Passive methods for their ease of use and low effort.

Security

While the accuracy differences between each mode are so small as to be immaterial there are other security considerations that may influence and your decision whether to implement Active or Passive Voice Biometrics:

- **Fraud Detection** – In addition to authenticating callers' identities, voice biometric systems can also be used to identify known fraudsters either by comparing the caller to a list of known bad actors or identifying that the same caller has purported to be multiple different customers in a period of time. These fraud detection methods are usually used when the caller has not been positively verified by Voice Biometrics but are only really applicable in Passive implementations, as fraudsters are likely to stay silent during the Active audio capture process and even if they didn't, the audio is too short to be used without a significant false alarm rate.

- **Recordings** – More correctly known as a Presentation Attack, this is a method by which a fraudster may attempt to circumvent a Voice Biometric system by playing back a recording of the genuine customer speaking or some use form of man-in-the-middle attack where the customer is socially engineered into saying the right things while conferenced into the authentication system. As a result of their more predictable nature, Active systems are far more vulnerable to this form of attack and while vendors have correspondingly developed a range of countermeasures which, if deployed, mitigate some of this risk but have their own false alarm rates and usability challenges reducing the end to end performance. While Passive systems are not immune to the same attack, the fact that the agent is having a conversation with the customer makes it significantly less predictable, and the agent is far more easily able to test that the caller is not a recording or that they are being duped.

- **Synthetic Voices** – While much has been written about increasing realism of deep fakes in video and audio, in practice they do not yet sufficiently represent all the variations in human speech that may be misrecognized by voice biometric systems. Again, vendors are constantly developing countermeasures to detect the tell-tale signs of fake audio and fraudsters are not yet sufficiently incentivized to deploy the significant effort (and duration of genuine customer audio) it takes to create even an approximate model. Both modes are equally vulnerable to this form of attack. However, for Active, the passphrase Is fixed, therefore it is much easier to capture a recording and this is therefore likely to be the most vulnerable mode in the short term.

Cost

There are also some differences in the cost and complexity of implementing each mode although this variation is reducing with time.

- **Technical Integration** – The acquisition of the required audio for Active voice biometrics is straightforward with most standards-based IVR platforms which can easily provide a short utterance to another application and process the results. In Passive implementations there is more complexity in acquiring an audio stream in real time, associating an identity with it (often from an agent's desktop application) processing it and returning the result to the correct agent in a timely fashion. The move to Voice Over IP platforms has, however, made this significantly easier but there is still a need for some customization as each end user's configuration and use case will be slightly different. The move towards cloud-based contact center platforms is increasing standardization of these configuration options, and we expect this will further reduce if not eliminate the amount of customization required in the future. On the agent desktop there will be development effort required for both Passive and Active modes but more for Passive as there is more direct engagement with the authentication and registration process. Most vendors now provide template user interfaces and standards-based interfaces that development teams will not struggle to use. For Active implementations development will be required to implement registration and authentication in the automated telephony application. On balance, we believe the differences largely cancel each other out but encourage you to conduct more detailed investigation of your chosen mode to avoid as soon as possible to avoid unexpected issues later during implementation.

- **Business Process Integration** – If implementing, as we recommend, agent led registration for Active systems, then in practice there is very little difference between Active and Passive voice biometrics in terms of the business processes that need to be designed and implemented. Both require agents to be able to offer the service, capture consent and complete enrollment. Both need supporting processes that handle customer queries about the service, delete enrollments on request and investigate anomalous results. The only additional requirement for Passive authentication is to handle occasional mismatches. Both approaches therefore are likely to require the same agent training, education and back office resources so this should not be a deciding factor in your decision between Active and Passive.

Passive Authentication in the IVR

Organizations should be aware that as the performance of Passive algorithms continues to improve we are approaching the point where it is possible to consider the use of Passive techniques to authenticate callers using the short utterances they provide to natural language and conversational call routing / self-service applications. Typically callers provide less than 3 seconds of speech to these services making it very challenging to provide high confidence results but with appropriate voice user interface design to increase average utterance length and algorithm improvement a significant proportion can now be authenticated this way. We expect this to be a major area of vendor focus in the next few years as it really is the best of both worlds although we expect that it is very unlikely that customers will be able to register this way and therefore it is still dependent on a Passive agent led registration process.

Conclusions and Next Steps

In conclusion, the decision of passive vs. active should be based on a thorough understanding of the current process context, weighing the usability, efficiency and security implications of each mode with a realistic assessment of the capabilities of agents, customers and contact center. Whichever mode an organization chooses Voice Biometrics will have a significant impact on the performance of the authentication process and has the potential to transform the experience of customers and agents alike. With hundreds of deployments and millions of registered users; organizations can rely on an established body of best practice to provide confidence of outcome whatever the mode of implementation chosen.

About SymNex

SymNex Consulting works with some of the most innovative and customer centric organizations to help them make the case for, design and implement transformational changes to the telephone welcome experience. Delivering dramatic improvements in the efficiency, security and convenience of these process through technology, pragmatism and behavioral understanding. symnexconsulting.com

LumenVox®

About LumenVox

LumenVox transforms customer communication. Our flexible and cost-effective technology enables you to create effortless, secure self-service and customer-agent interactions. We provide a complete suite of speech and authentication technology to make customer relations faster, stronger and safer than ever before. Our expertise is extensive— we support a multitude of applications for voice biometrics, inclusive of passive and active authentication for fraud detection. And we do it all by putting you and your customers first.

Interested in finding out more about this product?

Learn More →

Contact

LVsales@lumenvox.com +1 (858) 707-7700

www.LumenVox.com