

# The Emergence of Voice Verification

by [Kush R. Wadhwa](#)

Copyright August, 2004 Faulkner Information Services. All rights reserved.

---

## Inside this report ...

[Executive Summary](#)  
[Projected Revenues by Application](#)  
[Projected Total Revenues](#)  
[Risks](#)  
[Recommendations](#)  
[Action Plan for CSOs](#)  
[Resource File](#)

## Executive Summary

[return to [top](#) of report]

At a recent briefing of financial analysts, Bill Gates spoke about Microsoft's plans for driving growth in its business. Included among the productivity enhancements to its annual \$11 billion Office business was the incorporation of voice recognition capabilities, placing this biometric technology plainly in the mainstream view, even if not in daily use.<sup>1</sup>

Voice verification technology is based upon the distinctive characteristics derived from spoken phrases. These characteristics are based on the physiology of the vocal tract combined with unique behavioral aspects of speaking to enable verification of the identity of the person who is speaking. This should not be (but sometimes is) confused with speech recognition, a technology that translates what a person says into information that can be understood by a computer-based system.

Traditionally, enrollment in a voice verification system has been text dependent. In order to utilize voice verification technology, the individual must speak a specific phrase or word, repeating it two or three times at enrollment. Later, during authentication, the user normally only needs to recite the same word or phrase once. The phrases used during enrollment must be sufficiently rich in vocal content and generally range from two to five seconds in length.

Recently, we have begun to see the emergence of text-independent enrollment and verification. With such systems, the users can enroll by saying anything, from a simple conversation, to reading a paragraph of random text (e.g., an article from a newspaper). With these types of systems, the requisite length of the text recited to enroll is much longer – at least 30-45 seconds of text. Text independent systems allow enrollment to take place more passively, for example, while a customer is interacting with a call center agent;

however, the technology is not widespread at this time, with initial pilot implementations and deployments just beginning in the commercial sector.

It is also important to note that voice systems are currently limited to 1:1 verification (am I who I say I am?) applications – the future for the technology may allow for 1:N matching for identification (who am I?) applications, but the technology has not yet arrived at this point, nor has the market made this demand as yet. In point in fact, voice verification technologies currently represent only a small share of the overall biometric marketplace, but as the biometric market is anticipated to grow at a rapid pace over that next five years, the total voice verification market should grow to nearly \$225 million from its current size of \$23 million.<sup>2</sup>

There are approximately 20 companies in the voice verification market, many of which offer voice verification as one of many voice- and speech-related solutions. A handful of established companies dominate this market segment, providing core voice verification extraction and matching technology, packaged software solutions, and custom integrated hardware and software solutions (e.g., Nuance, Persay, ScanSoft, T-NETIX, VoiceVault). Currently, many of the leading voice verification solutions are developed by speech recognition companies and the natural synergies between these two industries will help drive revenue growth.

The following table illustrates how global voice verification revenues (in USD million) are expected to break down for various applications.<sup>3</sup>

**Table 1. Projected Revenues by Application**

[return to [top](#) of report]

| Application                       | 2003<br>\$(m) | 2004<br>\$(m) | 2005<br>\$(m) | 2006<br>\$(m) | 2007<br>\$(m) | 2008<br>\$(m) |
|-----------------------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Civil ID                          | 0.0           | 0.0           | 0.0           | 0.0           | 0.0           | 0.0           |
| Surveillance                      | 0.0           | 0.2           | 0.5           | 1.1           | 2.1           | 3.1           |
| PC / Enterprise Network<br>Access | 1.4           | 4.4           | 8.8           | 12.8          | 16.7          | 24.0          |
| Criminal ID                       | 0.0           | 0.0           | 0.0           | 0.0           | 0.0           | 0.0           |
| Retail / ATM / POS                | 0.7           | 2.0           | 4.4           | 9.1           | 15.6          | 20.7          |
| e-Commerce / Telephony            | 17.7          | 26.4          | 37.8          | 48.3          | 68.8          | 79.9          |
| Access Control / Attendance       | 0.0           | 0.0           | 4.1           | 5.8           | 7.5           | 9.4           |
| Device Access                     | 3.2           | 12.9          | 20.5          | 37.3          | 60.5          | 88.5          |

Copyright © 2004 International Biometric Group

The solutions offered do not require any specialized or proprietary acquisition devices, and everyone with a telephone is a potential user of the technology for voice verification. In

addition, these solutions leverage existing processes, already familiar to their users to both enroll and later authenticate within the system. For example, an existing process to speak an account number or personal data can be used to generate an enrollment template, and the same process can be used later to perform the verification. On the other hand, implementations where speech would need to be introduced as a new process do not represent a strong option, albeit some vendors are seeking innovative approaches to leverage the technology. One example is a new type of credit card being developed by Beepcard that requires successful voice verification against an on-board enrollment template before emitting an authorizing signal to the credit card provider.

The applications responsible for most of the growth in voice verification technology are in the commercial sector, enabling financial services account access and customer authentication for service calls. The driving force behind such deployments is the desire to reduce call center costs. Even a moderately accurate voice verification solution capable of biometrically authenticating 80% - 90% of users, while routing 10% - 20% of callers through standard authentication processes, can significantly reduce call center costs and ensure that operator screening efforts are directed at the most suspect callers.

Another example of ROI-driven deployments of voice verification is password resetting. For example, when users need to reset their passwords or PINs, traditionally they have had to call a help desk operator and verify their identity manually, typically by presenting some form of personal information. The cost to organizations for password reset has been estimated to be from \$200 to \$300 a year per user. Voice verification technology allows the automation of this process by positively verifying authorized users quickly, and then allowing them to change or reset their passwords themselves. As with caller authentication, users whose identity is deemed uncertain should always be redirected to a live operator.

Commercial applications such as these along with public-sector challenge-response implementations (e.g., house arrest and probation-related authentication) are expected to encompass the largest implementations for voice verification. In point of fact, all of these solutions commonly combine voice verification and speech recognition technologies, where spoken information is used to both authenticate identity and retrieve personal data.

Voice verification is used in conjunction with speech recognition products to provide automated access to accounts, ensuring (1) that the proper account is retrieved; and (2) that an authorized individual is accessing that account. As with all behavioral biometrics, voice verification can be used in combination with a "secret" value, such as the last four digits of a social security number, to increase security.

For example, Australia's Centrelink, the agency that provides personal details and information to welfare recipients, is currently exploring how to utilize voice verification to prevent estranged spouses from accessing their ex-partners' details. An imposter in such circumstances may well know some "secret" information such as a mother's maiden name, but would need to also be capable of defeating the biometric system.<sup>4</sup>

The expansion of telephony-based services will also drive an increase in deployments of voice verification, as consumers look for more granular access to information and transactions from land and mobile devices. Institutions will need to strike a balance between securing personal information and providing access to information. Voice verification can provide an extra level of security, enabling expanded services with increased security.

According to International Biometric Group's "Biometric Market Industry Report 2004-2008," the market for voice verification technology is expected to grow rapidly as enterprises and mobile telephony providers incorporate the core technology within their operations and

devices. 2003 revenues of \$23.0m are expected to grow to \$76.2m by 2005 and \$224.6m by 2008. While this technology is expected to comprise less than 5% of overall biometric revenues by the end of this period, the technology is well positioned for growth since it does not face very strong competition from alternative biometrics. With the exception of fingerprint sensors embedded in mobile phones, there are no competing biometric solutions for the applications in which voice verification technology is commonly deployed.

**Table 2. Projected Total Revenues**

[return to [top](#) of report]

|                                                               | 2003 | 2004 | 2005 | 2006  | 2007  | 2008  |
|---------------------------------------------------------------|------|------|------|-------|-------|-------|
| Projected Total Revenues (\$m)                                | 23.0 | 45.9 | 76.2 | 114.4 | 169.2 | 224.6 |
| Percentage of Total Revenues Attributed to Voice Verification | 3.2  | 3.8  | 4.1  | 4.3   | 4.6   | 4.8   |

Copyright © 2004 International Biometric Group

## Risks

[return to [top](#) of report]

Real-world testing has shown voice verification systems to be accurate under ideal conditions,<sup>5</sup> typically more so than how the technology is generally perceived. Nonetheless, there are clearly variations in performance between different vendors, and the conditions must be ideal. For example, if a user enrolls from a landline phone, but tries to verify on a mobile phone, there can be a distinct drop in performance, even with perfect reception. Background noise can also negatively affect performance, and as with all biometric technologies, the quality of the enrollment is key to optimal results. A longer enrollment utterance typically generates better performance, but there are diminishing returns after a certain point, as well as usability issues that arise if the enrollment process is too long. Further, there can be drops in performance over time, suggesting that a user's enrollment template should be updated after a period of time. Some systems address this problem through habituation, where the enrollment data is further refined after each successful verification.

Other major issues, which may present challenges to the effective implementation of voice verification applications include questionable accuracy, lack of wide scale deployment, and demographic variations.

### 1. Questionable Accuracy

The technology needs to offer a level of accuracy that assures confidence in the deployer and its customers. Customers are likely to assume that the systems prevent imposters from accessing their accounts, but if false non-match rates are high, and an inordinate number of calls must be redirected out the system to a live operator for intervention, the customer is likely to become frustrated and customer satisfaction levels, in addition to call center costs,

will suffer. The technology needs to be able to find a balance between security and convenience to be truly accepted.

## 2. Lack of Wide Scale Deployment

While the technology has potential to be broadly used, mainly because of the accessibility of acquisition devices (i.e., telephones), voice verification solutions have not been widely deployed in real-world applications. Because of this, potential deployers may not be fully convinced of the technology's accuracy and scalability and expansion to larger-scale deployments are expected to be incremental.

## 3. Demographic Variations

Not all systems perform equally well for all demographic groups. Gender, accents, and in some situations age can impact algorithm performance. Furthermore, certain individuals have "weak" voice prints more easily broken into than others, and certain individuals are more capable of breaking in against random accounts than others.

# Recommendations

[return to [top](#) of report]

In contemplating implementation of voice verification systems, potential deployers should consider the following key trends influencing how this technology may be used to address the overall goals (cost reduction, increased security, etc.) of their implementation in determining whether it is the "best" biometric for the application.

### 1. Evaluate ROI for Voice Verification

In a commercial environment, the ROI for voice verification systems can and should be evaluated. For call center cost avoidance, ROI can be readily calculated, although an organization's tolerances for false non-matching should be considered, as this will drive the percentage of calls redirected to live operators and ultimately impact the ROI.

### 2. Implement Where Voice Is Familiar

Use of voice verification will likely only be successful in environments where voice is already a part of an existing process. For example, it is not expected that PC-based voice verification will be widespread, except where users are already accustomed to speaking to their computer. On the other hand, the implementation of voice verification for access to mobile phones is a good fit. Voice transmissions may need to be classified as trusted or non-trusted, and to ensure this, voice verification logic will be built into mobile phone chipsets in the future. Note that fingerprint-based biometric firms are also targeting the use of mobile devices (telephones, PDAs) as an appropriate environment for deployment of sensors, and may dilute the market for voice vendors.

### 3. Watch for New Text-Independent Solutions

New applications, such as call monitoring, may be contemplated as text-independent solutions emerge.

**Table 3. Action Plan for CSOs**

[return to [top](#) of report]

| <b>Action</b>                       | <b>Purpose</b>                                                                                                                                                                                                                   |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Evaluate ROI for Voice Verification | To ensure voice verification makes economic sense.<br><br>To gain organizational support.<br><br>To provide a baseline for evaluating voice verification success.<br><br>To help calculate return on security investment (ROSI). |
| Implement Where Voice Is Familiar   | To facilitate the introduction of voice verification technology.<br><br>To reduce any reservations concerning implementation.                                                                                                    |

## Resource File

[return to [top](#) of report]

Anovea: <http://www.anovea.com/>  
Biometrics Consortium: <http://www.biometrics.org/>  
Find Biometrics: <http://www.findbiometrics.com/>  
Graphco: <http://www.graphcoholdingscorp.com/>  
International Biometric Group: <http://www.biometricgroup.com/>  
Microsoft: <http://www.microsoft.com/>  
Nuance: <http://www.nuance.com/>  
Persay: <http://www.persay.com/>  
Scansoft: <http://www.scansoft.com/>  
Speech Technology magazine: <http://www.speechtechmag.com/>  
Vocent: <http://www.vocent.com/>  
Voice Trust: <http://www.voicetrust.com/>  
Voicevault: <http://www.voicevault.com/>

## References

<sup>1</sup> S. Lohr. "Pursuing Growth, Microsoft Steps Up Patent Chase." New York Times. July 30, 2004.

<sup>2</sup> Biometric Market Industry Report, 2004 - 2008. International Biometric Group.

<sup>3</sup> ibid.

<sup>4</sup> "Voice verification move to stop welfare fraud." smh.com.au. June 4, 2003.

<sup>5</sup> "Comparative Biometric Testing, Rounds II – V." International Biometric Group.

## About the Author

[return to [top](#) of report]

Mr. Wadhwa is Director of Europe, Middle East and Asia with International Biometric Group. He performs strategy consulting for firms within the biometric industry, provides advisory services to private and public-sector deployers of biometrics on a global basis, with particular emphasis on border security, privacy and aviation and airport security. He is frequently published in industry and technology periodicals on biometrics-related topics, and regularly speaks on issues related to biometrics and smart cards. Recent presentations have included audiences such as the UK Biometric Working Group, the Teletrust Biometric Working Group, the Winter Biometrics Summit and the SecuTech Expo Taipei. Mr. Wadhwa can be reached at [kwadhwa@biometricgroup.com](mailto:kwadhwa@biometricgroup.com).

---

Site content copyright 2004, [Faulkner](#) Information Services. All rights reserved.

[Return to Security Management Practices Home](#)