

Dragon[®] speech recognition

Driving productivity and cost efficiency

Table of contents

- 2 Introduction**
- 2 Speed document turnaround using traditional transcription**
- 3 How does speech recognition work?**
- 4 Navigate applications by voice**
- 5 Automate complex business processes**
- 6 Increase productivity for mobile workers**
- 7 Managing a desktop deployment**
- 8 Justifying the expense**

Introduction

Government processes are driven largely by information – from field reports and case histories to requests for services and regulatory compliance paperwork. What if there was a faster, easier way to complete intensive documentation? The ability to automate repetitive or complex document creation tasks presents enormous opportunities for enhancing productivity, increasing efficiency and reducing costs.

Speech recognition technology can enable government agencies to dramatically reduce the time it takes to create documents, navigate complex information systems, or complete any other time consuming tasks that require staff to type information into a computer. Consider how much faster organizations could create and process reports if employees could dictate them instead of typing them. Imagine field reports—complete with accurate and detailed descriptions—created entirely by voice in a fraction of the time.

Speech solutions help to make Microsoft Windows applications faster and easier to use—whether employees are at their desks or in the field. In this way, speech recognition software can be used to boost employee productivity, provide accessibility for disabled workers, and protect employees from repetitive stress injuries. Government agencies can also deploy speech solutions more broadly to speed document turnaround, automate form completion, reduce transcription costs, and streamline repetitive workflows—without having to change existing business processes or existing information systems.

This white paper discusses some of the most pressing productivity issues facing today's government organizations and how speech recognition tools can help address both individual and enterprisewide challenges.

Speed document turnaround using traditional transcription

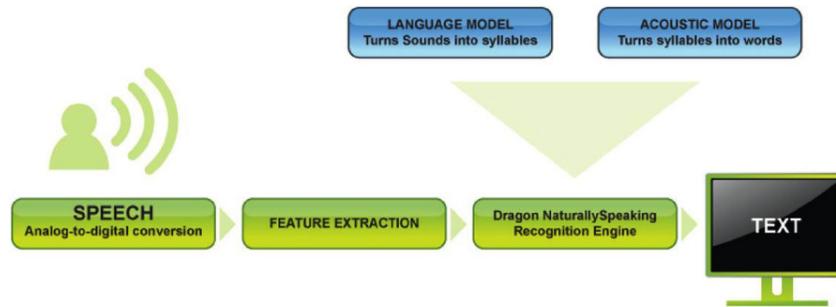
In recent years, government agencies have been tasked with handling a rapidly growing number of cases. At the same time, forms and reports—covering everything from education and social services to law enforcement and corrections—now require even more documentation for auditing and accountability purposes. Government employees are flooded with data that must be entered into the agency's information system. Too often, these employees are taking notes, writing reports, or filling in forms by hand and then entering that same information into an online document or application.

Resource-constrained agencies need an affordable solution that helps employees manage workloads more efficiently while improving the document management process. Speech recognition tools empower staff to create comprehensive reports, interviews, and other documents quickly and accurately.

Speech recognition delivers accurate dictation results three times faster than typing, enabling government workers to keep pace with an ever-increasing workload. For this reason, the FBI selected Dragon for its agents and professional support staff. Dragon offers the ability to enter and edit interview notes, reports, email messages, and other documents—all by voice—at rates of up to 160 words per minute.

What if there was a faster, easier way to complete intensive documentation? Consider how much faster agencies could create and process reports if employees could dictate them instead of typing them.

How does speech recognition work?



Desktop dictation allows government agencies to:

- Create and manage documents, emails, and forms faster than ever before
- Quickly access information in enterprise applications and Web sites
- More easily navigate and control desktop applications

Desktop speech recognition software products use the human voice as the main interface between the user and the computer. While relatively simple to use, speech recognition software is a highly sophisticated technology that leverages “language modeling” to recognize and differentiate among the millions of human utterances that make up any language. Using statistical models, speech recognition programs analyze an incoming stream of sound and interpret those sounds as commands and dictation. This process of interpretation is called speech recognition, and its success is measured by the percentage of correct interpretations.

Speech recognition systems create a voice profile for each user of the system that contains information about the unique characteristics of each person’s voice along with a customized set of words, known as a vocabulary, and user-specific information including software settings and personalized voice commands. When users create and train their user profile, they start with a standard set of models and then customize them for the way they speak (acoustic model) and the way they use words (vocabulary and associated language model). This approach accommodates users with varying accents and speech patterns. The software employs the customized user profile to guess the words spoken. Every time an individual uses the speech software and corrects his recognition errors, the software updates his user profile to enable better recognition accuracy over time.

Speech recognition works with the applications you already have: With appropriate speech recognition tools, users can simply talk to their computer and their words instantly appear in the full Microsoft Office Suite, as well as Microsoft Internet Explorer, Corel WordPerfect, Lotus Notes, and virtually all other Windows-based applications. Users can also create, navigate, send, and respond to email by voice using popular programs like Microsoft Outlook or Lotus Notes.

From creating documents and email messages to writing detailed reports and completing routine forms, speech speeds document turnaround to enable faster response times. For example, an officer in Kabul, Afghanistan, was required to submit daily and weekly reports to his agency and senior staff. Using Dragon speech recognition, he was able to cut the report generation time from three hours of typing to 45 minutes of dictation. Saving more than two hours per day from typing freed him to focus more on his core duties.

Navigate applications by voice

Speech recognition software enables users to command and control the computer desktop by voice. Virtually any menu item or dialog box can be controlled hands-free. Users can edit and format their work, launch applications and open files, or cut-and-paste documents. In other words, speech recognition helps to speed up routine tasks on the PC.

Many applications can be easier to use and more effective when deployed in conjunction with speech recognition. Searches, queries, and form filling are all faster to perform by voice than keyboarding. Document management, document assembly/automation, and database software programs are all highly conducive to control by speech. Tasks such as text and data entry can be completed by voice in most programs without any customization. Other functions can easily be performed using macro voice commands (see below).

Managing email: Since managing email takes up an increasing amount of everyone's workday, speech recognition software can be used to create, navigate, send, and respond to email by voice using popular programs like Microsoft Outlook or Lotus Notes. Dragon Professional makes email use even faster and easier with the Dragon Voice Shortcuts for Email feature: A user can simply say, "Send email to Case Worker XYZ" and Dragon will launch the user's email program, create a new email and put the name(s) the user said into the "To:" box.

In addition, some speech recognition programs contain text-to-speech technology that allows users to have their email documents read aloud, which enables them to complete other tasks while listening to their email.

Work on the Web by voice: Today's speech recognition programs make it easier than ever for government workers to search the Web, access information, and navigate Web pages. This includes not only the public Internet, but also private Intranets and other HTML interfaces.

Most speech recognition programs allow users to speak standard commands that prompt the computer to perform an action. For example, the user says, "start WordPerfect," and the PC launches WordPerfect. Transferring this concept to information searches, Dragon Professional enables users to collapse the multiple tasks associated with Web and desktop searches into single voice commands. Users can search the Web or their desktop with a simple voice command such as, "Search the Web for global warming articles," or "Find email about the Robinson report."

Ideal for employees with physical challenges of repetitive strain

injuries: In most cases, speech recognition is used in conjunction with other input devices including the keyboard and mouse.

However, users can leverage advanced speech recognition products, such as Dragon Professional, to control 100 percent of their computing environment, making this technology ideal for employees with physical challenges, repetitive strain injuries or other reasons to operate information systems completely hands-free.

Because Dragon Professional is Section 508 compliant, it can help organizations meet disability management needs. It meets the US government's standards for ensuring software applications are accessible to people with disabilities.

Where is speech recognition in use in government agencies today?

- Field reporting
 - Caseworkers
 - Legal departments
 - Accessibility (RSI, paralysis, dyslexia, etc.)
 - Law enforcement
 - Large document creation
 - Automating legacy systems
 - Physical notes transcription
 - General office productivity
-

"I discovered that this exciting career came with a tremendous amount of PAPERWORK! Through the years, I continued to drown in the ocean of very detailed reports. Unlike those television detectives, the true life detectives must submit their investigations in report fashion. Dragon allowed me to dictate reports and finally begin to see the light at the end of the tunnel."

- Tennessee police sergeant

Up to 99 percent accurate: Speed means nothing without accuracy. Today's speech recognition tools can be up to 99 percent accurate right out of the box. But every government agency uses particular terminology, acronyms, or other vocabulary unique to its area of specialization or its regional location. These unique terms are frequently used in forms, reports, and other documents.

To boost accuracy – and productivity – advanced speech recognition tools enable agencies to:

- Add new words or customize the vocabulary with employee names, acronyms, and specialized terminology frequently used in their particular department or area of public service
- Delete entries that could cause acoustic ambiguity
- Analyze an individual's written documents to update the user profile based on writing style and words used

If a government organization uses particular names or terminology with a high degree of frequency, customizing the vocabulary across the entire enterprise can increase recognition accuracy so users spend less time correcting errors. Enterprise speech recognition products enable sharing of customized vocabularies among all employees—or a subset of employees—eliminating the need to update each user profile separately.

Automate complex business processes

Multi-step processes consolidated into a single voice command: An enterprise deployment of speech recognition allows federal, state, or local agencies to go beyond dictation to accomplish routine tasks more quickly and efficiently. One approach is to cut down on the number of steps it takes to complete a given task—without changing established business processes.

Every day, case workers, police officers and others perform repetitive tasks involving multiple applications, mouse clicks, and keyboard entries. With speech recognition, these multiple tasks can be consolidated and executed with a single voice command.

Government employees spend a considerable amount of time creating reports and processing forms—many of which share standard elements. As a result, many employees find themselves entering the same information into these documents time and time again. These repetitive, time-consuming processes can be a huge productivity drain on overworked government agencies. By simplifying multi-step processes that employees perform dozens of times a day, workflow automation can deliver significant productivity gains, especially when multiplied across hundreds of workers enterprise wide.

Boilerplate commands: Users can create text blocks—including commonly used phrases, paragraphs and even graphics—and insert them into documents or emails using a single voice command for faster, easier document creation. These custom commands can also contain variable “voice fields.” In this way, a single voice command creates a complete document that can be customized by navigating through each field to fill in variable information (e.g., the name of a client or a fee on a standard client letter or contract template).

Government agencies can add unique names and terms in custom vocabularies to increase recognition accuracy rates. Users spend less time correcting errors. For example, if many of a state prosecutor's cases are heard by a judge named John Shaffer, the prosecutor's office may want to consider deleting the other common spellings of the name Shaffer (Schaeffer, Schaefer, Shafer, etc). This way, Dragon will choose the intended spelling for more accurate dictation of court reports and other legal documents.

Macro commands: Repetitive tasks, such as data entry or form filling, can be sped up by using speech. In many cases, users who are unfamiliar with complex software programs are more comfortable “telling” the computer what to do rather than trying to master the interface. Macros can be created to enable users to go from field to field by voice, or to perform a sequence of keystrokes or mouse movements. Even skilled typists are often slower at entering numbers and letters. Spreadsheets can be created and edited using voice commands. Accounting and time billing software can also be controlled by voice.

For example, government workers—or more typically the IT departments that support them—can use Microsoft Visual Basic to build a single voice command that saves a contract, emails it to a standard list of recipients at the client company, and prints out a hard copy at both corporate headquarters and the appropriate regional sales office—all with a single spoken command such as “Complete Contract.”

Repetitive tasks are far less time-consuming when they’re automated with simple voice commands. By simplifying multi-step processes that staff performs dozens of times a day, workflow automation can deliver significant productivity gains, especially when multiplied across hundreds of workers enterprise wide.

Increase productivity for mobile workers

Many government employees—case workers, patrol officers, health and building inspectors, probation officers, and more—spend their workdays on the road, traveling from location to location. In many cases, these employees have seen their workloads increase sharply as budgets shrink and demand for services increases.

With speech recognition tools, government workers can take care of business from virtually any location—whether they’re in the office, at a field location, or on the road. Simply dictate into a handheld device for automatic transcription when the user returns to his PC. Dictate notes, reports, and other documents—safely and accurately—to complete paperwork in a more timely and efficient manner. This mobile productivity tool ensures that time on the road or off site does not mean a paperwork pile-up upon return to the office.

Field agents can also dictate reports, forms, and other documents by using a Bluetooth wireless microphone as an input device to get more work done without being tethered to a PC. A user can search the Internet and navigate Web pages by simply speaking URLs, links, or voice commands for fast, efficient information access—such as when an officer needs to access information from his patrol car.

Roaming users: Dragon enables government employees—whether they’re building inspectors, case workers, state troopers, or any other mobile worker—to use and retrieve their user profile from a desktop at the office, on a laptop in their vehicle, or on a home PC logged into the central network. These profiles are available to end users anywhere on the network via mapped drive, UNC Path or http.

System administrators centrally manage individual user profiles—complete with macros and customized vocabularies – and perform centralized updates and synchronization. By configuring Dragon so that user profiles reside on the

Automate processes with macro commands: Repetitive tasks take far less time when they’re automated with simple voice commands. With Dragon NaturallySpeaking Professional, a set of abbreviated instructions can be given to the computer to complete repeatable tasks that would normally take multiple keyboard combinations and mouse clicks.

central network, government IT organizations can ensure that the software increases its accuracy each time an employee uses it—regardless of where their jobs take them.

With user profiles located on the central network, administrators can run the Acoustic and Language Optimizer for all users (vs. expecting individual users to run the tool locally). Since no other activity on the machine is possible as the process uses 100% of the CPU resources, individuals may skip this tool. But running the Acoustic and Language Model Optimization regularly is critical to increase the overall accuracy of the system. In an enterprise deployment, the Optimizer can be centrally scheduled by the administrator to run during off-hours.

Managing a desktop deployment

A successful speech recognition deployment requires careful attention to user expectations, training, and customization. Some organizations choose to manage their own speech recognition installation, customization, and training, but most prefer to outsource this work to the software manufacturer, a system integrator, or speech recognition value added reseller (VAR).

In a distributed government enterprise, IT professionals are responsible for administration of software tools deployed across many locations on a variety of computing devices. Imagine if an Oracle database deployed across the enterprise lacked centralized administration capabilities. The result would be chaos and frustration as IT scrambled to manage individual users for each system change.

To simplify administration, lower costs, and ensure data security, government organizations need enterprise-strength tools and applications based on industry standards that can be installed and managed from a central network location. The same requirements apply to desktop dictation solutions. Products intended for personal use simply don't offer the network capabilities required for enterprise deployments.

Network administration: Installing and managing a desktop dictation solution from a central network location enables system administrators to:

- Create and manage installations and user profiles over a network
- Distribute customized vocabularies and commands automatically
- Control settings
- Restrict access to specific features on a user-by-user basis
- Automatically synchronize updates and changes via a variety of communication protocols
- Perform system backups

Advanced speech recognition systems offer administrative tools that enable enterprise users to share custom vocabularies and macros across multiple users. For example, the administrator can use vocabulary enhancements made by one individual in her own profile and let other users benefit from these enhancements. Updates of shared vocabularies can be pushed to multiple end-users automatically. This eliminates the time-consuming task of entering new words and pronunciations one at a time on each end-user's machine.

Keeping up with caseloads: Jim Coursin, a caseworker for children at Clinton County Protective Services in New York, travels around the county conducting hundreds of interviews per year. The handwritten notes taken at all those interviews must be entered into the agency's Windows-based case management system. Even given Jim's impressive typing speed of 60 words per minute, inputting all those notes was an extremely time-consuming process. A few years ago, Coursin started using Dragon. Now he dictates his notes into a digital recorder and Dragon automatically transcribes the text. Thanks to Dragon, Jim not only saves about an hour a day in text input, but he is able to provide more detailed and complete notes about each interview.

A centralized network allows multiple users to discover best practices and quickly and easily share them across the user community. It also fosters the creation of an Intranet “Dragon Information Center” to post specific demos, tips & tricks, FAQs, sample word lists and commands, a “suggestion box” and more.

Customization: An investment in vocabulary customization can also deliver big payoffs. By customizing the system’s vocabulary at the start of an engagement, enterprises can obtain remarkably accurate recognition results from the first day of deployment. In addition, the speech experts at Nuance or members of the Value-Added Reseller community can work with an organization up front to understand its processes and identify opportunities for increasing efficiency, then create and deliver vocabularies as well as sets of custom commands and macro commands that can speed the execution of repetitive, multistep tasks to deliver big productivity benefits when shared across enterprise users. While text and step-by-step macros can be created with no knowledge of programming, more complex macros require advanced scripting using Microsoft Visual Basic.

Some organizations have IT departments that can create such macros for Dragon users, after some instruction on the Dragon’s specific functions; alternatively, organizations may wish to enlist assistance to develop some or all of the initial commands, and a few “super-users” are taught.

Justifying the expense

Enterprise deployments of Dragon speech recognition consist of several components:

- Client software
- Professional Services (planning, installation, customization, training and support)
- Audio peripherals (headsets, digital recorders, wireless microphones)
- Enterprise resources (Server & storage resources, back end system integration, end-user support, data and profile maintenance)

In most cases, enterprises that purchase Dragon NaturallySpeaking realize improved productivity and return on investment (ROI) almost immediately. What makes this rapid ROI possible?

- **It’s easy to use.** For a sophisticated tool, Dragon is remarkably easy to use—allowing most users to be up and running in less than 15 minutes—leading to high adoption rates with minimal training and support costs.
- **It saves time.** Dragon enables users to create documents and emails three times faster than typing.
- **It’s accurate.** With recognition accuracy rates of up to 99 percent, Dragon allows users to quickly create detailed and accurate reports—without any spelling errors.
- **It’s fast.** Macros automate and streamline repetitive manual processes for productivity increases of up to 300 percent.

A probation and parole officer suggests that Dragon has tripled his department’s output, cutting report creation time from three days to one day. “Business as usual” consisted of dictating reports on a recorder then submitting the recorded reports to a secretary to have them typed. Due to

Products intended for personal use simply don’t offer the network capabilities required for enterprise deployments.

“With Dragon, I feel I can do the job of essentially two to three police officers in that I’m not tied up doing paperwork at a police department somewhere,” stated Gary Bonzo, a police officer in Clearwater, Florida. “Instead I’m out on the road doing patrols and I’m able to respond to calls faster.”

the volume of reports to be typed, it would take three days before the reports would be typed and returned. “In my line of work, the public, my employers, attorneys and the courts expect concise, accurate and timely reports,” he explains. Senior staff members at the agency were convinced that Dragon could be used to enhance the department’s overall productivity.

Personnel-related costs: An ROI evaluation for Dragon begins with a review of the intrinsic loss of personnel working on transcription tasks instead of on primary tasks, based on the average hourly salary. But consider, too, the cost of other personnel-related items:

- Estimated annual cost of noncompliance with American Disabilities Act in computer operations (legal fees, law suit awards/settlements, lost business opportunities, etc.)
- Cost of computer-related RSI and similar claims
- Estimated annual loss of personnel productivity from RSI
- When the benefits of Dragon are multiplied across dozens or even hundreds of users enterprise wide, the cost savings and productivity gains add up quickly.

About Nuance Communications, Inc.

Nuance Communications is reinventing the relationship between people and technology. Through its voice and language offerings, the company is creating a more human conversation with the many systems, devices, electronics, apps and services around us. Every day, millions of people and thousands of businesses experience Nuance through intelligent systems that can listen, understand, learn and adapt to your life and your work. For more information, please visit nuance.com.
